

A;Accession: S26496
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 65-100 <DEV>
A;Cross-references: EMBL:X64023
A;Accession: S26497
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 3-97 <DE2>
A;Cross-references: EMBL:X64022
C;Genetics:
A;Gene: p9Ka
A;Introns: 47/3
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: calcium binding; EF hand
F;50-82/Domain: calmodulin repeat homology <EF2>

Query Match 91.6%; Score 488; DB 2; Length 101;
Best Local Similarity 91.1%; Pred. No. 1.2e-39;
Matches 92; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
QY 1 MACPLEKALDVMTSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
DB 1 MARPLEALDVMTSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
QY 61 NLDSNRDNEVDFOEYCVFLSCIAMMNCNEFFEGPDK 101
DB 61 NLDSNRDNEVDFOEYCVFLSCIAMMNCNEFFEGPDK 101

RESULT 5
A30129
S-100 protein, lung - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 01-Sep-1995
R;Glenney Jr., J.R.; Kindy, M.S.; Zokas, L.
J. Cell Biol. 108, 569-578, 1989
A;Title: Isolation of a new member of the S100 protein family: amino acid sequence, tissue
A;Reference number: A30129; MUID:89139574; PMID:2521861
A;Accession: A30129
A;Status: not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-97 <GLE>
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: calcium binding; EF hand
F;50-82/Domain: calmodulin repeat homology <EF2>

Query Match 63.0%; Score 336; DB 2; Length 97;
Best Local Similarity 64.9%; Pred. No. 3.6e-25;
Matches 61; Conservative 17; Mismatches 16; Indels 0; Gaps 0;
QY 1 MACPLEKALDVMTSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
DB 1 MSSPLEQALAVMTSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
QY 61 NLDSNRDNEVDFOEYCVFLSCIAMMNCNEFFEGP 94
DB 61 NLDSNRDNEVDFOEYCVFLSCIAMMNCNEFFEGP 94

RESULT 6
A41988
S-100 calcium-binding protein A2 - human
N;Alternate names: calcium-binding protein Can19; S-100 calcium-binding protein L
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 29-Aug-1997
C;Accession: A41988
R;Lee, S.W.; Tomasetto, C.; Swishhelm, K.; Keyomarsi, K.; Sager, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 2504-2508, 1992
A;Title: Down-regulation of a member of the S100 gene family in mammary carcinoma cells
A;Reference number: A41988; MUID:92196147; PMID:1372446
A;Accession: A41988

A;Molecule type: mRNA
A;Residues: 1-98 <LEE>
A;Cross-references: GB:M87068; NID:gl79896
A;Experimental source: mammary carcinoma cells
A;Note: sequence extracted from NCBI backbone (NCBIN:88561, NCBIP:88562)
C;Genetics:
A;Gene: GDB:S100A2; CAN19; S100L
A;Cross-references: GDB:211118; OMIM:176993
A;Map position: 1q21-1q21
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: calcium binding; EF hand
F;8-42/Domain: calmodulin repeat homology <EF1>
F;51-83/Domain: calmodulin repeat homology <EF2>

Query Match 61.5%; Score 328; DB 2; Length 98;
Best Local Similarity 62.5%; Pred. No. 2.1e-24;
Matches 60; Conservative 17; Mismatches 19; Indels 0; Gaps 0;
QY 1 MACPLEKALDVMTSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
DB 2 MCSLEQALAVMTSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 61
QY 61 NLDSNRDNEVDFOEYCVFLSCIAMMNCNEFFEGPDK 96
DB 62 NLDSNRDNEVDFOEYCVFLSCIAMMNCNEFFEGPDK 97

RESULT 7
B28363
calcyclin - rat
N;Alternate names: prolactin receptor-associated protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Sep-1989 #sequence_revision 30-Sep-1989 #text_change 16-Jul-1999
C;Accession: B28363
R;Murphy, L.C.; Murphy, L.J.; Tsuyuki, D.; Duckworth, M.L.; Shiu, R.P.C.
J. Biol. Chem. 263, 2397-2401, 1988
A;Title: Cloning and characterization of a cDNA encoding a highly conserved, putative cal
A;Reference number: A92705; MUID:88115387; PMID:2448309
A;Accession: B28363
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-90 <MUR>
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: calcium binding; EF hand
F;48-80/Domain: calmodulin repeat homology <EF2>

Query Match 46.9%; Score 250; DB 2; Length 90;
Best Local Similarity 51.1%; Pred. No. 5.2e-17;
Matches 47; Conservative 20; Mismatches 23; Indels 2; Gaps 1;
QY 1 MACPLEKALDVMTSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
DB 1 MACPLDQALGLAVMTSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 58
QY 61 NLDSNRDNEVDFOEYCVFLSCIAMMNCNEFFEG 92
DB 59 DLDRNRDNEVDFOEYCVFLSCIAMMNCNEFFEG 90

RESULT 8
BCHUY
calcyclin - human
N;Alternate names: growth factor-inducible protein 2A9; prolactin receptor-associated pro
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1988 #sequence_revision 31-Dec-1992 #text_change 22-Jun-1999
C;Accession: A28363; A25645; A30356; PGL222
R;Murphy, L.C.; Murphy, L.J.; Tsuyuki, D.; Duckworth, M.L.; Shiu, R.P.C.
J. Biol. Chem. 263, 2397-2401, 1988
A;Title: Cloning and characterization of a cDNA encoding a highly conserved, putative cal
A;Reference number: A92705; MUID:88115387; PMID:2448309
A;Accession: A28363
A;Molecule type: mRNA
A;Residues: 1-90 <MUR>

R;Baudier, J.; Gerard, D.
Biochemistry 22, 3360-3369, 1983
A;Title: Ions binding to S100 proteins: structural changes induced by calcium and zinc
A;Reference number: A90471; MUID:84000339; PMID:16615778
A;Contents: annotation; metal ion-binding properties
R;Okazaki, K.; Obata, N.H.; Inoue, S.; Hidaka, H.
Biochem. J. 306, 551-555, 1995
A;Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in glia
A;Reference number: S54343; MUID:95194333; PMID:7887910
A;Accession: S54346
A;Molecule type: protein
A;Residues: 24-33 <OKA>
C;Comment: The S-100 protein is composed of two related polypeptide chains, alpha and beta
brain proteins, S-100 is also found in a variety of other tissues.
C;Comment: S-100 is an intracellular protein that weakly binds calcium. It binds zinc ve
es, with different affinities, exist for both ions on each monomer. Physiological conce
calcium-binding sites.
C;Superfamily: S-100 protein; calmodulin repeat homology
F;2-94/Product: S-100 protein alpha chain #status predicted <MAT>
F;7-41/Domain: calmodulin repeat homology <EF1>
F;50-82/Domain: calmodulin repeat homology <EF2>
F;2/Modified site: blocked amino end (Gly) (in mature form) (probably acetylated) #statu
F;20,23,25,28,33/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted
F;63,65,67,69,74/Binding site: calcium (Asp, Asn, Asp, Glu, Glu) #status predicted

Query Match 46.0%; Score 245; DB 1; Length 94;
Best Local Similarity 51.1%; Pred. No. 1.6e-16;
Matches 46; Conservative 14; Mismatches 30; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTFHKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
DB 1 MGSELETAMETLLNVFHAHSGEGDKYKLSKKELKELLQTELSGFLDAQKDADAVDKVMK 60

QY 61 NLDNSRDNEVDFOEYCVFLSCIAMMCNEFF 90
DB 61 ELDENGEGEVDFOEYVVLVAALTVCNNEFF 90

RESULT 11
S27011
calcylin - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 16-Jun-2000
A;Accession: S27011; S20264; S16116
R;Ando, I.; Watanabe, M.; Akatsuka, H.; Tokumitsu, H.; Hidaka, H.
FEBS Lett. 314, 109-113, 1992
A;Title: Site-directed mutation makes rabbit calcylin dimer.
A;Reference number: S27011; MUID:93093152; PMID:1459239
A;Accession: S27011
A;Molecule type: mRNA
A;Residues: 1-90 <AND>
A;Cross-references: EMBL:D10895; NID:G217743; PIDN:BAA01707.1; PID:G217744
A;Experimental source: lung
R;Tokumitsu, H.; Kobayashi, R.; Hidaka, H.
Arch. Biochem. Biophys. 291, 401, 1991
A;Title: Corrigendum. A calcium-binding protein from rabbit lung cytosol identified as b
A;Reference number: S20264; MUID:92061074; PMID:1952954
A;Accession: S20264
A;Molecule type: protein
A;Residues: 23-30;35-46;55-89 <OK>
R;Tokumitsu, H.; Kobayashi, R.; Hidaka, H.
Arch. Biochem. Biophys. 288, 202-207, 1991
A;Title: A calcium-binding protein from rabbit lung cytosol identified as the product of
A;Reference number: S16116; MUID:91378440; PMID:1898017
A;Accession: S16116
A;Molecule type: protein
A;Residues: 23-27,'G',29-30;35-42,'G',44-46;55-81,'V',83-89 <OK>
C;Superfamily: S-100 protein; calmodulin repeat homology
F;Keywords: calcium binding; EF hand; monomer
F;7-41/Domain: calmodulin repeat homology <EF1>

F;48-80/Domain: calmodulin repeat homology <EF2>

Query Match 45.8%; Score 244; DB 1; Length 90;
Best Local Similarity 51.1%; Pred. No. 1.9e-16;
Matches 47; Conservative 18; Mismatches 25; Indels 2; Gaps 1;

QY 1 MACPLEKALDVMTFHKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
DB 1 MASPLDQAIGLLIGIFHKYSGEGDKHLSKKELKELLQTEL--IGSKLQDABIVKMD 58

QY 61 NLDNSRDNEVDFOEYCVFLSCIAMMCNEFF 92
DB 59 DLDRNKDOEWNFOEYITFLGALAMINAEALKG 90

RESULT 12
BCHUIA
N;Alternate names: S-100 calcium-binding protein A1 (S100A1)
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence revision 23-May-1997 #text_change 22-Jun-1999
A;Accession: A44470; S16740
R;Engelkamp, D.; Schafer, B.W.; Erne, P.; Heizmann, C.W.
Biochemistry 31, 10258-10264, 1992
A;Title: S100 alpha, CAPL, and CACY: molecular cloning and expression analysis of three c
A;Reference number: A44470; MUID:93041710; PMID:1384693
A;Accession: A44470
A;Molecule type: mRNA
A;Residues: 1-94 <ENG>
A;Cross-references: EMBL:X58079; NID:G36175; PIDN:CAA41107.1; PID:G36176
A;Experimental source: heart
A;Note: sequence extracted from NCBI backbone (NCBI:116494)
C;Comment: This protein binds p53, tubulin and many other proteins at physiological conce
C;Comment: S-100 is an intracellular protein that binds calcium. It binds zinc more tight
different affinities exist for both ions on each monomer. Physiological concentrations c
nding sites.
C;Comment: Although predominant among the water-soluble brain proteins, S-100 is also fo
C;Genetics:
A;Gene: GDB:S100A1; S100A
A;Cross-references: GDB:126839; OMIM:176940
A;Map position: 1q21-1q21
C;Complex: homodimer; heterodimer with S-100 protein beta chain (see PIR:BCHUIB)
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: brain; calcium binding; EF hand; heterodimer; homodimer; zinc
F;7-41/Domain: calmodulin repeat homology <EF1>
F;50-82/Domain: calmodulin repeat homology <EF2>

Query Match 45.8%; Score 244; DB 1; Length 94;
Best Local Similarity 51.1%; Pred. No. 2e-16;
Matches 46; Conservative 14; Mismatches 30; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTFHKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
DB 1 MGSELETAMETLLNVFHAHSGEGDKYKLSKKELKELLQTELSGFLDAQKDADAVDKVMK 60

QY 61 NLDNSRDNEVDFOEYCVFLSCIAMMCNEFF 90
DB 61 ELDENGEGEVDFOEYVVLVAALTVCNNEFF 90

RESULT 13
S35985
S-100 protein alpha chain - weatherfish
C;Species: Misgurnus fossilis (weatherfish)
C;Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 10-Sep-1999
A;Accession: S35985
R;Ivanenkov, V.V.; Gerke, V.; Minin, A.A.; Plessmann, U.; Weber, K.
Mech. Dev. 42, 151-158, 1993
A;Title: Transduction of Ca(2+) signals upon fertilization of eggs; identification of an
A;Reference number: S35985; MUID:94031845; PMID:8217841
A;Accession: S35985
A;Molecule type: protein
A;Residues: 1-95 <IVA>

RESULT 15
S20330
calvasculin - bovine (fragments)
C;Species: Bos primigenius taurus (cattle)
C;Date: 19-Mar-1997 #sequence_revision 29-May-1998 #text_change 29-Sep-1999
C;Accession: S20330
R;Watanabe, Y.; Kobayashi, R.; Ishikawa, T.; Hidaka, H.
Arch. Biochem. Biophys. 292, 563-569, 1992
A;Title: Isolation and characterization of a calcium-binding protein derived from mRNA
A;Reference number: S20330; MUID:92117670; PMID:1731618
A;Accession: S20330
A;Molecule type: protein
A;Residues: 1-10;11-55 <WAT>

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:42:57 ; Search time 21.6387 Seconds
(without alignments)
189.536 Million cell updates/sec

Title: US-10-067-618-6

Perfect score: 65

Sequence: 1 LPSFLGKRTDEAA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_thc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_xvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	%	Description
1	65	100.0	101	6 Q9TV56
2	55	84.6	55	6 Q9TSB1
3	44	67.7	521	10 Q9SNZ5
4	41	63.1	741	16 Q8FNS3
5	40	61.5	111	16 Q8D110
6	40	61.5	130	10 Q04815
7	40	61.5	131	16 Q82QY3
8	40	61.5	354	9 Q8SCT8
9	39	60.0	554	10 Q9SI51
10	39	60.0	707	16 Q88CE5
11	39	60.0	972	10 Q88AG3
12	38	58.5	211	10 Q9FY27
13	38	58.5	234	16 Q8ZAO8
14	38	58.5	246	16 Q8CWI0
15	38	58.5	530	5 Q9USK0
16	38	58.5	530	5 Q7YUK5

17	38	58.5	2109	12	Q98776	Q98776 vesicular s
18	38	58.5	2109	12	Q86125	Q86125 vesicular s
19	38	58.5	2109	12	Q8B010	Q8B010 vesicular s
20	38	58.5	2109	12	Q8B0H5	Q8B0H5 vesicular s
21	38	58.5	2109	12	Q8B0H0	Q8B0H0 vesicular s
22	37	56.9	66	16	Q8RG13	Q8RG13 fusobacteri
23	37	56.9	161	10	Q40935	Q40935 pseudotsuga
24	37	56.9	161	10	Q40936	Q40936 pseudotsuga
25	37	56.9	161	10	Q8LSD5	Q8LSD5 pseudotsuga
26	37	56.9	233	10	Q8S2N4	Q8S2N4 oryza sativ
27	37	56.9	263	16	Q7V6W2	Q7V6W2 prochloroco
28	37	56.9	286	16	Q8FMK5	Q8FMK5 corynebacte
29	37	56.9	293	15	Q9DX93	Q9DX93 human immun
30	37	56.9	294	15	Q9DX92	Q9DX92 human immun
31	37	56.9	308	10	Q9SNT7	Q9SNT7 oryza sativ
32	37	56.9	331	16	Q8PGA7	Q8PGA7 xanthomonas
33	37	56.9	380	16	Q8DUR2	Q8DUR2 streptococc
34	37	56.9	411	16	Q8EQP5	Q8EQP5 oceanobacil
35	37	56.9	431	16	Q9A856	Q9A856 caulobacter
36	37	56.9	500	2	Q9VNN6	Q9VNN6 escherichia
37	37	56.9	500	16	Q9EYD2	Q9EYD2 schizosacch
38	37	56.9	566	3	Q9P7G6	Q9P7G6 schizosacch
39	37	56.9	722	5	Q810U5	Q810U5 aplysia kur
40	37	56.9	735	5	Q9NDE8	Q9NDE8 aplysia cal
41	37	56.9	784	10	Q9FLG1	Q9FLG1 arabidopsis
42	37	56.9	797	10	Q7XVN7	Q7XVN7 oryza sativ
43	37	56.9	986	16	Q8YS66	Q8YS66 anabaena sp
44	37	56.9	1280	4	Q86SQ6	Q86SQ6 homo sapien
45	36.5	56.2	261	2	Q83VG7	Q83VG7 neisseria p

ALIGNMENTS

RESULT 1

Q9TV56 PRELIMINARY; PRT; 101 AA.
AC Q9TV56
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Metastasin.
GN MTS1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Medin-Durby; TISSUE=Kidney;
RA Miyamori H., Hasegawa K., Kim K., Sato H.;
RT "Expression of metastasin associated mts1 gene is co-induced with
RT membrane type-1 matrix metalloproteinase (MT1-MMP) during oncogenic
RT transformation and tubular formation of madin darby canine kidney
RT (MDCK) epithelial cells."
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AB031064; BAA83419.1; -.
DR HSSP; P30801; 1A03.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
SQ SEQUENCE 101 AA; 11847 MW; 5535387DB7577DF0 CRC64;

Query Match 100.0%; Score 65; DB 6; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.00025;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RN	[3]
RA	SEQUENCE FROM N.A.
RL	EU Arabidopsis sequencing project;
DR	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
EMBL; AL035540; CAB37512.1; -	
PIR; TOS684; TOS684; -	
InterPro; IPR007942; PEARLI-4.	
Pfam; PF05278; PEARLI-4; 1.	
SQ SEQUENCE 521 AA; 58777 MW; 9502A09624B59432 CRC64;	
Query Match 67.7%; Score 44; DB 10; Length 521;	
Best Local Similarity 66.7%; Pred. No. 12;	
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps	
OY 1 LPSFLGKRTDEA 12	
Db 67 LPSVYGKRMDES 78	
RESULT 4	
Q8FNS3 PRELIMINARY; PRT; 741 AA.	
ID Q8FNS3	
AC Q8FNS3;	
DT 01-MAR-2003 (TREMBlrel. 23, Created)	
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)	
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)	
DE Putative eukaryotic-type serine/threonine protein kinase.	
CN CE2070.	
OS Corynebacterium efficiens.	
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;	
OC Corynebacteriaceae; Corynebacterium.	
NCBI_TaxID=152794;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;	
RA Kawabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,	
Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,	
Usuda Y., Sugimoto S.;	
"The entire genomic sequence of Corynebacterium efficiens YS-314.";	
Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.	
EMBL; AF005221; BAC18880.1; -	
GO; GO:0005224; F:ATP binding; IEA.	
GO; GO:0008658; F:penicillin binding; IEA.	
GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.	
GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.	
GO; GO:0006468; P:protein amino acid phosphorylation; IEA.	
InterPro; IPR005543; PASTA.	
InterPro; IPR000719; Prot_kinase.	
InterPro; IPR002290; Ser_thr_kinase.	
InterPro; IPR008271; Ser_thr_pkin_AS.	
InterPro; IPR001245; Tyr_pkinase.	
Pfam; PF03793; PASTA; 3.	
Pfam; PF00069; pkinase; 1.	
ProDom; PD000001; Prot_kinase; 1.	
SMART; SM00220; S_TKC; 1.	
SMART; SM00219; TyrcKc; 1.	
PROSITE; PS00111; PROTEIN KINASE DOM; 1.	
PROSITE; PS00108; PROTEIN KINASE ST; 1.	
KW Kinase; Serine/threonine-protein kinase; Complete proteome.	
SQ SEQUENCE 741 AA; 78804 MW; 655FB3058197C2B9 CRC64;	
Query Match 63.1%; Score 41; DB 16; Length 741;	
Best Local Similarity 58.3%; Pred. No. 63;	
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps	
OY 1 LPSFLGKRTDEA 12	
Db 686 VPSLLGRVVDA 697	
RESULT 5	
Q8DII0	

```

ID Q8D110 PRELIMINARY; PRT; 111 AA.
AC Q8D110;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein.
GN Y0458.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611 (2002).
DR EMBL; AB013646; AAM94047.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016462; F:pyrophosphatase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR008162; Pyrophosphatase.
DR InterPro; IPR001130; TatD_DNase.
DR Pfam; PF01026; TatD_DNase; 1.
DR PROSITE; PS00387; PFAS; 1.
DR PROSITE; PS01091; TATD_3; 1.
KW Hypothetical protein.
SQ SEQUENCE 111 AA; 12657 MW; 56EDAA13C56C411F CRC64;

Query Match 61.5%; Score 40; DB 16; Length 111;
Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 2 PSFLGKRTDEAA 13
Db |::|::|::|::|
93 FNLGQKTDENA 104

RESULT 6
O04815 PRELIMINARY; PRT; 130 AA.
AC O04815;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Hypothetical protein (Fragment).
OS Sporobolus stapfianus (Resurrection grass).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Chloridoideae; Eragrostideae; Sporobolus.
OX NCBI_TaxID=56623;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=leaf;
RX Blomstedt C.K., Gianello R.D., Neale A.D., Hamill J.D., Gaff D.F.;
RT "Isolation and characterization of cDNAs associated with the onset of
RT desiccation tolerance in the resurrection grass Sporobolus
RT stapfianus.";
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y10779; CAA71751.1; -.
DR InterPro; IPR004238; LEA.
DR Pfam; PF02987; LEA; 2.
KW Hypothetical protein.
FT NON_TER 130
FT NON_TER 130
SQ SEQUENCE 130 AA; 13735 MW; A2D7C05BF28AFE6A CRC64;

Query Match 61.5%; Score 40; DB 10; Length 130;

```

```

Best Local Similarity 70.0%; Pred. No. 16;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 SFLGKRTDEA 12
Db |||||::|::|
15 SFLGQKTDQA 24

RESULT 7
Q82QY3 PRELIMINARY; PRT; 131 AA.
AC Q82QY3;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein.
GN SAV362.
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=21477403; PubMed=11572948;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonce T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680 / ATCC 31267 / NCIMB 12804 / NRRL 8165;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531 (2003).
DR EMBL; AP005022; BAC68071.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 131 AA; 13413 MW; B2928813E42EF4F2 CRC64;

Query Match 61.5%; Score 40; DB 16; Length 131;
Best Local Similarity 69.2%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
Db |||||::|::|
114 LPSFLGPPPTAA 126

RESULT 8
Q8SCY8 PRELIMINARY; PRT; 354 AA.
AC Q8SCY8;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE PHK2174.
OS Pseudomonas phage phik2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OX NCBI_TaxID=169683;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21914557; PubMed=11916376;
RX Mesyanzhinov V.V., Robben J., Grymonprez B., Kostyuchenko V.A.,
RA Bourkaltseva M.V., Sykilla N.N., Krylov V.N., Volckaert G.;
RT "The genome of bacteriophage phik2 of Pseudomonas aeruginosa.";
RL J. Mol. Biol. 317:1-19 (2002).
RN [2]

```

```

RP SEQUENCE FROM N.A.
RA Mesyanzhinov V.V., Robben J., Grymonprez B., Kostyuchenko V.A.,
RA Bourkaltseva M.V., Syklinda N.N., Krylov V.V., Volkkaert G.;
RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF399011; AAL83075.1; -.
SQ SEQUENCE 354 AA; 39093 MW; F81242EF7F0B3F32 CRC64;

Query Match 61.5%; Score 40; DB 9; Length 354;
Best Local Similarity 61.5%; Pred. No. 46;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
DB 6 LPSFIGTGKDEVA 18

RESULT 9
Q9SI51 PRELIMINARY; PRT; 554 AA.
AC Q9SI51;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Disease resistance protein (TIR-NBS-LRR class), putative
DE (Fragment).
GN AT2G14080.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC007197; RAD25848.3; -.
DR PIR; H84513; H84513.
DR InterPro; IPR001611; LRR.
DR Pfam; PF00560; LRR; 7.
FT NON TER 1
SQ SEQUENCE 554 AA; 62216 MW; BD92DA9CDF5B6ACC CRC64;

Query Match 60.0%; Score 39; DB 10; Length 554;
Best Local Similarity 70.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPSFLGKRTD 10
DB 196 LPSFIGKATN 205

RESULT 10
Q8CE5 PRELIMINARY; PRT; 707 AA.
AC Q8CE5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN Pf2336.
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,

```

```

RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzapple E., Scanlan D., Tran K., Moazzez A.,
RA Uterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016793; AAN70801.1; -.
DR TIGR; PP5236; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 707 AA; 78798 MW; B3C05A55B0D9903E CRC64;

Query Match 60.0%; Score 39; DB 16; Length 707;
Best Local Similarity 70.0%; Pred. No. 1.4e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 LPSFLGKRTD 10
DB 212 LPRYLGRVD 221

RESULT 11
Q8S8G3 PRELIMINARY; PRT; 972 AA.
ID Q8S8G3;
AC Q8S8G3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Disease resistance protein (TIR-NBS-LRR class), putative
DE (Fragment).
GN AT2G14080.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1];
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
RA Fraser C.M., Venter J.C.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006528; AAM15274.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; P:ATP binding; IEA.
DR GO; GO:0003793; P:defense/immunity protein activity; IEA.
DR GO; GO:0000166; P:nucleotide binding; IEA.
DR GO; GO:0004888; P:transmembrane receptor activity; IEA.
DR GO; GO:0006915; P:apoptosis; IEA.
DR GO; GO:0006952; P:defense response; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR000767; Disease_resist.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR002182; NB-ARC.
DR InterPro; IPR000157; TIR.
DR Pfam; PF00560; LRR; 6.
DR Pfam; PF00931; NB-ARC; 1.
DR Pfam; PF01582; TIR; 1.
DR PRINTS; PR00364; DISEASERSIST.
DR SMART; SM00382; AAA; 1.
DR SMART; SM00255; TIR; 1.
DR PROSITE; PS0104; TIR; 1.
KW ATP-binding.
FT NON TER 972
SQ SEQUENCE 972 AA; 110045 MW; DB0CAB6399507496 CRC64;

Query Match 60.0%; Score 39; DB 10; Length 972;
Best Local Similarity 70.0%; Pred. No. 2e+02;

```

Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPSFLGKRTD 10
 |||||
 Db 857 LPSFIGKATN 866
 |||||

RESULT 12
 Q9FYZ7 PRELIMINARY; PRT; 211 AA.
 ID Q9FYZ7
 AC Q9FYZ7
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
 DE CTR (Fragment)
 OS Nicotiana tabacum (Common tobacco)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC lamids; Solanales; Solanaceae; Nicotiana.
 OC NCBI_TaxID=4097;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Ky57;
 RA Terajima Y., Satoh S.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF247567; AAG00418.1; -;
 FT NON TER 1 1
 FT NON TER 211 211
 SQ SEQUENCE 211 AA; 22848 MW; EB11EB199E64A1B4 CRC64;

Query Match 58.5%; Score 38; DB 10; Length 211;
 Best Local Similarity 58.3%; Pred. No. 64;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 PSFLGKRTDEAA 13
 |||||
 Db 25 ENFLGPADESA 36
 |||||

RESULT 13
 Q8ZAQ8 PRELIMINARY; PRT; 234 AA.
 ID Q8ZAQ8
 AC Q8ZAQ8
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Endonuclease V (EC 3.1.-.-).
 GN NFI OR YPO3733.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OC NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX PARKHILL J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdono-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Feltwell T., Haulin N., Holroyd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moulé S., Oyston F.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrall B.G.;
 RT "Genome sequence of Yersinia pestis, the causative agent of plague."
 RL Nature 413:523-527(2001).
 DR EMBL; AJ414158; CAC93201.1; -;
 DR PIR; AE0454; AE0454
 DR GO; GO:0016787; F:hydrolase activity; IEA.
 DR InterPro; IPR007581; Endonuc V.
 DR Pfam; PF04493; Endonuc V; 1.
 KW Hydrolase; Complete proteome.
 SQ SEQUENCE 234 AA; 26244 MW; 36287B1C274653A5 CRC64;

Query Match 58.5%; Score 38; DB 16; Length 234;
 Best Local Similarity 77.8%; Pred. No. 71;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PSFLGKRTD 10
 |||||
 Db 224 PDFLGKRRD 232
 |||||

RESULT 14
 Q8CWIO PRELIMINARY; PRT; 246 AA.
 ID Q8CWIO
 AC Q8CWIO
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Endonuclease V (Deoxyinosine 3'endonuclease).
 GN NFI OR Y0497.
 OS Yersinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Yersinia.
 OC NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KIMS / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of Yersinia pestis KIM."
 RL J. Bacteriol. 184:4601-4611(2002).
 DR EMBL; AE013650; AAM84086.1; -;
 DR InterPro; IPR007581; Endonuc V.
 DR Pfam; PF04493; Endonuc V; 1.
 SQ SEQUENCE 246 AA; 27620 MW; 1CA0BB7F95B9FE4 CRC64;

Query Match 58.5%; Score 38; DB 16; Length 246;
 Best Local Similarity 77.8%; Pred. No. 75;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PSFLGKRTD 10
 |||||
 Db 236 PDFLGKRRD 244
 |||||

RESULT 15
 Q9USKO PRELIMINARY; PRT; 530 AA.
 ID Q9USKO
 AC Q9USKO
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Trypanosoma brucei.
 OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OC NCBI_TaxID=5691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ILTari;
 RX MEDLINE=21273154; PubMed=11377739;
 RA Maier A.G., Webb H., Ding M., Bremser M., Carrington M., Clayton C.;
 RT "The coatomer of Trypanosoma brucei."
 RL Mol. Biochem. Parasitol. 115:55-61(2001).
 DR EMBL; AJ250726; CAB60082.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 530 AA; 60091 MW; BCCA7B5191DB5E7 CRC64;

Query Match 58.5%; Score 38; DB 5; Length 530;
 Best Local Similarity 61.5%; Pred. No. 1.6e+02;
 Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
||| :|
Db 93 LPEFLRKRNFESA 105

Search completed: September 9, 2004, 10:58:20
Job time : 23.6387 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:32:01 ; Search time 31.5484 Seconds
(without alignments)
134.340 Million cell updates/sec

Title: US-10-067-618-7

Perfect score: 85
Sequence: 1 NEFFEGFPDKQPRKK 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: geneseqp1980s.*
2: geneseqp1990s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	15	2	AAR80457 Human mts
2	85	100.0	15	3	AB37436 Human mts
3	85	100.0	15	6	ABU08521 Human mts
4	85	100.0	101	2	AAR20560 Human mts
5	85	100.0	101	2	AAR80453 Human mts
6	85	100.0	101	3	AB45534 Human S10
7	85	100.0	101	3	AB37432 Human mts
8	85	100.0	101	4	AB72386 Human mts
9	85	100.0	101	6	ABU08513 Human mts
10	85	100.0	101	7	ADD14157 Human src
11	74	87.1	101	1	AB72387 Murine Mt
12	43	50.6	758	5	ABP99368 Arabidops
13	42	49.4	101	2	AAV12102 Human 5'
14	42	49.4	141	5	AB89552 Human pol
15	42	49.4	759	4	AG91772 C glutami
16	42	49.4	864	7	ADC30960 Human nov
17	42	49.4	1205	7	ADC10162 Human NOV
18	42	49.4	1234	4	AAE07171 Human Hul
19	42	49.4	1268	4	ABB11796 Human sec
20	41	48.2	97	2	AAR27058 Sequence
21	41	48.2	97	2	AAR26406 Sequence
22	41	48.2	97	3	AB45532 Human S10
23	41	48.2	97	5	ABG96408 Human ova
24	41	48.2	97	6	ABR92148 Human cer
25	41	48.2	97	7	ADB70348 S100 calc

ALIGNMENTS

RESULT 1
AAR80457
ID AAR80457 standard; peptide; 15 AA.

XX AAR80457;
XX
XX 27-DEC-1995 (first entry)
XX Human mts-1 (87-101) peptide 4.
XX Metastasis; cancer; mts-1 gene; tumour; therapy; antigen; antibody.
XX Synthetic.
XX WO9520656-A1.
XX
XX 03-AUG-1995.
XX
XX 31-JAN-1995; 95WO-US001214.
XX
XX 31-JAN-1994; 94US-00190560.
XX
XX (RESE) RESEARCH CORP TECHNOLOGIES INC.

XX Zain S, Lukanidin E;
XX WPI; 1995-275441/36.
XX Nucleic acid encoding human mts-1, antigenic fragments and antibodies -
XX useful for diagnosis of malignant cancer and metastatic potential of
XX tumour cells.
XX Disclosure; Page 24; 124pp; English.
XX
XX Antigenic peptides 1-4 (given in AAR80454-57) of human mts-1 protein
XX (AAR80453) were used to raise polyclonal and monoclonal antibodies that
XX detect the presence of mts-1 in tissue samples, esp. metastatic cells.
XX Peptides 1, 3 and 4 are unique to mts-1, while peptide 2 generates
XX antibodies reactive with many calcium binding proteins
XX
XX Sequence 15 AA;

Query Match 100.0%; Score 85; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
|||||

Ad559560 Human Pro
Ad559556 Human Pro
Abp71986 Human sta
Abu56414 Lung canc
Abu56415 Lung canc
Abu56412 Lung canc
Abu18599 Protein e
Aag10126 Arabidops
Aag10125 Arabidops
Aag10124 Arabidops
Aaw80937 Murine le
Aaw80939 Murine le
Ade25555 Human LLP
Aaw80940 Human kid
Aaw80938 Human kid
Aaw80942 Human hea
Aaw80944 Murine ki
Aay41708 Human PRO
Aab19579 Human PRO

```

Db      1 NEFFEGFPDKQPRKK 15

RESULT 2
AAB37436
ID AAB37436 standard; peptide; 15 AA.
XX
AC AAB37436;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human mts-1 peptide fragment #4.
XX
KW Human; mts-1; cytostatic; cancer; metastasis.
XX
OS Homo sapiens.
XX
PN WO200064475-A1.
XX
PD 02-NOV-2000.
XX
PF 20-APR-2000; 2000WO-US011006.
XX
PR 23-APR-1999; 99US-00298625.
XX
PA (RESE ) RESEARCH CORP TECHNOLOGIES INC.
XX
PI Lukanidin E;
XX
DR WPI; 2000-687266/67.
XX
PT Treating cancer and preventing metastasis comprises administration of an
XX anti-mts-1 protein antibody or antisense oligonucleotide.
XX
PS Example 9; Page 28; 155pp; English.
XX
CC The present invention relates to methods for treating cancer and
CC preventing metastases, comprising the administration of a composition
CC directed against the mts-1 protein (see AAC68131-C68132 and AAB37432).
CC Mts-1 protein is a calcium-binding protein, and is thought to have a role
CC in myoepithelial cell differentiation. The present sequence is a peptide
CC fragment of the human mts-1 protein. This peptide was used to generate
CC antibodies against mts-1 protein, which can be used to detect mts-1
CC protein in clinical specimens
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 85; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
   |||||
Db 1 NEFFEGFPDKQPRKK 15
   |||||

RESULT 3
ABU08521
ID ABU08521 standard; peptide; 15 AA.
XX
AC ABU08521;
XX
DT 27-MAY-2003 (first entry)
XX
DE Human mts-1 protein, antigenic peptide #4.
XX
KW Human; mts-1; metastatic cancer; p53; cytostatic; gene therapy; antigen.
XX
OS Homo sapiens.
XX
PN US2002172680-A1.
XX
PD 21-NOV-2002.

XX 29-APR-2002; 2002US-00135152.
XX
XX 09-JUL-1990; 90US-00550600.
XX
PR 25-NOV-1992; 92US-00981455.
XX
PR 31-JAN-1994; 94US-00190560.
XX
PR 06-JUN-1995; 95US-00468942.
XX
PR 23-APR-1999; 99US-00298635.
XX
PA (LUKA/) LUKANIDIN E.
XX
PI Lukanidin E;
XX
XX WPI; 2003-328422/31.
XX
PT Treating cancer by administering a reagent directed against the mts-1
XX protein or an oligonucleotide capable of binding to mts-1 mRNA.
XX
PS Example 9; Page 19; 64pp; English.
XX
CC The invention relates to treating cancer comprising administering a
CC reagent directed against the mts-1 (not defined) protein or an
CC oligonucleotide capable of binding to mts-1 mRNA. Also include are a
CC method of inactivating, destroying or nullifying a mts-1 protein or cells
CC and a method of inhibiting metastasis in a cancerous cell. Experiments
CC showed that mts-1 inhibited the phosphorylation of full-size p53 and the
CC C-terminal protein fragment by PKC. Addition of the same concentrations
CC of mts-1 to the PKC reaction mixture did not affect the phosphorylation
CC of the N-terminal and DNA-binding domains of p53. The method is useful
CC for treating cancers associated with the mts-1 gene. The present sequence
CC represents a human mts-1 protein antigenic peptide
XX
SQ Sequence 15 AA;

Query Match 100.0%; Score 85; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 4.1e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
   |||||
Db 1 NEFFEGFPDKQPRKK 15
   |||||

RESULT 4
AAR20560
ID AAR20560 standard; protein; 101 AA.
XX
XX AAR20560;
XX
DT 25-MAR-2003 (revised)
DT 14-MAY-1992 (first entry)
XX
DE Human mts protein.
XX
KW Metastatic cancer; antibodies; mouse; lung; liver; kidney; thyroid;
KW breast cancer; cell growth.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Protein 1..101
FT Peptide /label= mts-1
FT Peptide 2..11
FT Peptide /note= "antigenic"
FT Peptide 22..37
FT Peptide /note= "antigenic; calcium binding domain"
FT Peptide 42..54
FT Peptide /note= "antigenic"
FT Peptide 87..101
FT Peptide /note= "antigenic"
XX
PN WO9200757-A.

```

PD 23-JAN-1992.
 XX
 PF 09-JUL-1990; 90US-00550600.
 XX
 PR 09-JUL-1990; 90US-00550600.
 XX
 PA (RESE) RESEARCH CORP TECHNOLOGIES INC.
 XX
 PI Zain S, Lukanidin E;
 XX
 DR WPI; 1992-056647/07.
 DR N-PSDB; AAQ20506.
 XX
 PT Metastatic cancer diagnosis by detection of mts-1 gene or protein - using
 PT antibody treatment of cancer and tumours of e.g. kidney, thyroid, lung
 PT and liver.
 XX
 PS Claim 9; Fig 2; 82pp; English.
 XX
 CC The sequence was deduced from the DNA sequence obtd. by screening a human
 CC cDNA library with mouse mts-1 cDNA probes. The antigenic Mts-1 peptides
 CC (see features) derived from the protein and anti- bodies raised to them
 CC are useful in the diagnosis of metastatic cancer, e.g. lung, kidney,
 CC thyroid or breast cancer. The peptide comprising the calcium binding site
 CC generates antibodies reactive with many members of the calcium binding
 CC protein family; the other three peptides are unique to mts-1 and generate
 CC antibodies specific only for this protein. Cell lines capable of
 CC expressing mts-1 are useful as model systems for in vitro and in vivo
 CC anti-metastasis drug screening. Pharmaceutical compns. contg. the mts-1
 CC protein or anti-cancer reagents may be used to promote cell growth, or
 CC for treating cancer, respectively. (Updated on 25-MAR-2003 to correct PA
 CC field.)
 XX
 SQ Sequence 101 AA;

Query Match 100.0%; Score 85; DB 2; Length 101;
 Best Local Similarity 100.0%; Pred. NO. 2.8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NEFEFGPDQPRKK 15
 |||||
 Db 87 NEFEFGPDQPRKK 101
 |||||

RESULT 5
 AAR80453 ID AAR80453 standard; protein; 101 AA.
 XX
 AC AAR80453;
 XX
 DT 27-DEC-1995 (first entry)
 XX
 DE Human mts-1 protein.
 XX
 KW Metastasis; cancer; mts-1 gene; tumour; therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9520656-A1.
 XX
 PD 03-AUG-1995.
 XX
 PF 31-JAN-1995; 95WO-US001214.
 XX
 PR 31-JAN-1994; 94US-00190560.
 XX
 PA (RESE) RESEARCH CORP TECHNOLOGIES INC.
 XX
 PI Zain S, Lukanidin E;
 XX
 DR WPI; 1995-275441/36.
 DR N-PSDB; AAQ99177.
 XX

PT Nucleic acid encoding human mts-1, antigenic fragments and antibodies -
 PT useful for diagnosis of malignant cancer and metastatic potential of
 XX tumour cells.
 XX
 PS Claim 12; Page 92; 124pp; English.
 XX
 CC A human cDNA library was constructed in lambda-gt10 using poly(A)+ RNA
 CC prep'd. from HeLa cells. The library was screened with a 32P- labeled
 CC mouse mts-1 vDNA probe. A clone was obtd. which comprised the full-length
 CC human mts-1 gene. The encoded protein is used for the diagnosis or
 CC therapy of cancer, and to raise antibodies
 XX
 SQ Sequence 101 AA;

Query Match 100.0%; Score 85; DB 2; Length 101;
 Best Local Similarity 100.0%; Pred. NO. 2.8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NEFEFGPDQPRKK 15
 |||||
 Db 87 NEFEFGPDQPRKK 101
 |||||

RESULT 6
 AAB45534 ID AAB45534 standard; protein; 101 AA.
 XX
 AC AAB45534;
 XX
 DT 22-FEB-2001 (first entry)
 XX
 DE Human S100A4 protein.
 XX
 KW S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;
 KW calcium-binding protein; calcium homeostasis; cardiac muscle;
 KW pumping capacity; myocardial cell; systolic calcium ion release;
 KW sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;
 KW valve defect.
 XX
 OS Homo sapiens.
 XX
 PN DE19915485-A1.
 XX
 PD 19-OCT-2000.
 XX
 PF 07-APR-1999; 99DE-01015485.
 XX
 PR 07-APR-1999; 99DE-01015485.
 XX
 PA (KATU/) KATUS H A.
 PA (REMP/) REMPPIS A.
 XX
 PI Katus HA, Remppis A;
 XX
 DR WPI; 2000-673510/66.
 DR N-PSDB; AAC81804.
 XX
 PT Composition containing S100 protein, corresponding nucleic acid or
 PT vector, useful for treating cardiomyopathy and cardiac insufficiency.
 XX
 PS Claim 35; Page 11; 36pp; German.
 XX
 CC This invention describes a novel composition for treating primary or
 CC secondary cardiomyopathy or cardiac insufficiency contains at least one
 CC S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or
 CC fragments, or a gene transfer vector containing (II), optionally
 CC formulated with auxiliaries and/or carriers. (I) are calcium-binding
 CC proteins involved in calcium homeostasis, so their overexpression in
 CC cardiac muscle will improve pumping capacity (and overall capacity) of
 CC the heart. In cultured myocardial cells they increase the contraction and
 CC relaxation rates associated with increased systolic calcium ion release
 CC from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are
 CC used to treat cardiomyopathy (CMP) where inherited or caused by

CC spontaneous mutations and ischemic CMP caused by arteriosclerosis,
 CC dilative CMP caused by toxic/infectious disease, cardiac disease caused
 CC by pulmonary and/or arterial hypertension, and structural disease caused
 CC by rhythm disorders or valve defects, generally any condition associated
 CC with reduced contractile force. Unlike calmodulin, which is expressed
 CC ubiquitously, (I) show tissue-specific expression and treat the
 CC underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac
 CC disease

XX Sequence 101 AA;

Query Match 100.0%; Score 85; DB 3; Length 101;
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDQPRKK 15
 DB 87 NEFFEGFPDQPRKK 101

RESULT 7

ABR37432
 ID AAB37432 standard; protein; 101 AA.

XX AAB37432;

DT 21-FEB-2001 (first entry)

XX Human mts-1.

XX Human; mts-1; cytostatic; cancer; metastasis.

OS Homo sapiens.

XX WO200064475-A1.

PD 02-NOV-2000.

XX 20-APR-2000; 2000WO-US011006.

XX 23-APR-1999; 99US-00298625.

XX (RESE) RESEARCH CORP TECHNOLOGIES INC.

XX Lukanidin E;

XX WPI; 2000-687266/67.

DR N-PSDB; AAC68131, AAC68132.

XX Treating cancer and preventing metastasis comprises administration of an
 PT anti-mts-1 protein antibody or antisense oligonucleotide.

XX Example 6; Fig 2; 155pp; English.

XX The present sequence is human mts-1. Mts-1 protein is a calcium-binding
 CC protein, and is thought to have a role in myoepithelial cell
 CC differentiation. The present invention relates to methods for treating
 CC cancer and preventing metastases, comprising the administration of a
 CC composition directed against the mts-1 protein

XX Sequence 101 AA;

Query Match 100.0%; Score 85; DB 3; Length 101;
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDQPRKK 15
 DB 87 NEFFEGFPDQPRKK 101

RESULT 8

AAB72386

ID AAB72386 standard; protein; 101 AA.

XX AAB72386;

DT 24-MAY-2001 (first entry)

XX Human Mts 1 protein amino acid sequence.

XX Mts 1; Mts1-del75; Mts1-4S; S100A4; neuronal cell growth; cardiac arrest;
 KW neuronal degeneration; Parkinson's disease; Alzheimer's disease;
 KW Down's syndrome; stroke; sciatic crush; spinal cord injury; human.

OS Homo sapiens.

XX WO200118043-A2.

PD 15-MAR-2001.

XX 07-SEP-2000; 2000WO-US024495.

XX 10-SEP-1999; 99US-00393433.

XX (PROL-) PROLIFIA INC.

XX Bock E, Lukanidin EM, Berezin V;

XX WPI; 2001-235188/24.

XX New isolated functional derivatives of Mts protein for stimulating
 PT neuronal cell growth to treat neurological conditions e.g. Parkinson's
 PT disease, Alzheimer's disease, Down's syndrome and stroke.

XX Disclosure; Page; 60pp; English.

XX This invention relates to functional derivatives of an Mts protein,
 CC particularly Mts1-del75, and Mts1-4S. Mts 1 (also known as S100A4) is a
 CC member of the S100 gene family. Mts1 proteins and their functional
 CC derivatives are used to stimulate the growth of neuronal cells and in the
 CC treatment of neurological conditions including those characterised by
 CC neuronal degeneration, death or injury such as Parkinson's disease,
 CC Alzheimer's disease, Down's syndrome, stroke, cardiac arrest, sciatic
 CC crush, spinal cord injury, injury to sensory neurons or degenerative
 CC disease of the retina. The present sequence represents the human Mts 1
 CC protein, from which the Mts proteins of the invention may be derived.
 CC NOTE: The present sequence is not shown in the specification but it has
 CC been derived from the sequence given in US patent 5801142 (referred to on
 CC page 10 of the specification)

XX Sequence 101 AA;

Query Match 100.0%; Score 85; DB 4; Length 101;
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDQPRKK 15
 DB 87 NEFFEGFPDQPRKK 101

RESULT 9

ABU08513

ID ABU08513 standard; protein; 101 AA.

XX AC ABU08513;

XX 27-MAY-2003 (first entry)

XX Human mts-1 protein.

XX Human; mts-1; metastatic cancer; p53; cytostatic; gene therapy.

XX Homo sapiens.

XX

PN US2002172680-A1.
 XX 21-NOV-2002.
 XX 29-APR-2002; 2002US-00135152.
 XX 09-JUL-1990; 90US-0050600.
 PR 25-NOV-1992; 92US-00981455.
 PR 31-JAN-1994; 94US-00190560.
 PR 06-JUN-1995; 95US-00468942.
 PR 23-APR-1999; 99US-00298635.
 XX (LUKA/) LUKANIDIN E.
 PA Lukanidin E;
 XX WPI; 2003-328422/31.
 DR N-PSDB; ABX93578.
 XX Treating cancer by administering a reagent directed against the mts-1
 PT protein or an oligonucleotide capable of binding to mts-1 mRNA.
 XX Example 6; Fig 2; 64pp; English.
 CC The invention relates to treating cancer comprising administering a
 CC reagent directed against the mts-1 (not defined) protein or an
 CC oligonucleotide capable of binding to mts-1 mRNA. Also include are a
 CC method of inactivating, destroying or nullifying a mts-1 protein or cells
 CC and a method of inhibiting metastasis in a cancerous cell. Experiments
 CC showed that mts-1 inhibited the phosphorylation of full-size p53 and the
 CC C-terminal protein fragment by PKC. Addition of the same concentrations
 CC of mts-1 to the PKC reaction mixture did not affect the phosphorylation
 CC of the N-terminal and DNA-binding domains of p53. The method is useful
 CC for treating cancers associated with the mts-1 gene. The present sequence
 CC represents human mts-1 protein
 XX Sequence 101 AA;
 SQ

Query Match 100.0%; Score 85; DB 6; Length 101;
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NEFFEGPDKQPRKK 15
 |||||
 Db 87 NEFFEGPDKQPRKK 101

RESULT 10
 ADD14157
 ID ADD14157 standard; protein; 101 AA.
 XX AC ADD14157;
 AC
 DT 01-JAN-2004 (first entry)
 XX
 DE Human src biomarker polypeptide SEQ ID NO:346.
 XX predictor set; protein tyrosine kinase activity modulator;
 KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
 KW gene therapy; drug sensitivity; genetic profile; cancer; human.
 XX Homo sapiens.
 OS
 XX WO2003062395-A2.
 FN
 XX 31-JUL-2003.
 PD
 XX 17-JAN-2003; 2003WO-US001981.
 PF
 XX 18-JAN-2002; 2002US-0350061P.
 PR
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.

PI Huang F, Fairchild CR, Lee FY, Shaw P;
 XX WPI; 2003-636735/60.
 DR N-PSDB; ADD14760.
 XX New polynucleotides and polypeptides for predicting the activity of
 PT compounds that interact with protein tyrosine kinases and/or protein
 PT tyrosine kinase pathways.
 XX Claim 10; SEQ ID NO 346; 139pp; English.
 XX The present invention describes a predictor set comprising a plurality of
 CC polynucleotides or polypeptides whose expression pattern is predictive of
 CC the response of cells to treatment with a compound that modulates protein
 CC tyrosine kinase activity or members of the protein tyrosine kinase
 CC pathway. Also described: (1) predicting whether a compound is capable of
 CC modulating the activity of cells, comprising obtaining a sample of cells,
 CC determining whether the cells express a plurality of markers, and
 CC correlating the expression of the markers to the compound's ability to
 CC modulate the activity of the cells; (2) a plurality of cell lines for
 CC identifying polynucleotides and polypeptides whose expression levels
 CC correlate with compound sensitivity or resistance of cells associated
 CC with a disease state; and (3) identifying polynucleotides and
 CC polypeptides that predict compound sensitivity or resistance of cells
 CC associated with a disease state, comprising subjecting the plurality of
 CC cell lines to one or more compounds, analyzing the expression pattern of
 CC a microarray of polynucleotides or polypeptides, and selecting
 CC polynucleotides or polypeptides that predict the sensitivity or
 CC resistance of cells associated with a disease state by using the
 CC expression pattern of the microarray. The polynucleotides and
 CC polypeptides have cytostatic activities, and can be used in gene therapy.
 CC The polynucleotides and polypeptides are useful in predicting the
 CC activity of compounds that interact with protein tyrosine kinases and/or
 CC protein tyrosine kinase pathways. These may be used in determining drug
 CC sensitivity in patients to allow the development of individualized
 CC genetic profiles which aid in treating diseases and disorders (e.g.
 CC cancer) based on patient response at a molecular level. The present
 CC sequence is used in the exemplification of the present invention.
 XX Sequence 101 AA;
 SQ

Query Match 100.0%; Score 85; DB 7; Length 101;
 Best Local Similarity 100.0%; Pred. No. 2.8e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NEFFEGPDKQPRKK 15
 |||||
 Db 87 NEFFEGPDKQPRKK 101

RESULT 11
 AAB72387
 ID AAB72387 standard; protein; 101 AA.
 XX AC AAB72387;
 AC
 DT 24-MAY-2001 (first entry)
 XX Murine Mts 1 protein amino acid sequence.
 DE
 XX Mts 1; Mts1-del75; Mts1-4S; S100A4; neuronal cell growth; cardiac arrest;
 KW neuronal degeneration; Parkinson's disease; Alzheimer's disease;
 KW Down's syndrome; stroke; sciatic crush; spinal cord injury; mouse.
 XX Mus sp.
 OS
 XX WO200118043-A2.
 FN
 XX 15-MAR-2001.
 PD
 XX 07-SEP-2000; 2000WO-US024495.
 PF
 XX 10-SEP-1999; 99US-00393433.
 PR

XX (PROL-) PROLIFIA INC.
 XX Bock E, Lukanidin EM, Berezin V;
 XX WPI; 2001-235188/24.
 XX New isolated functional derivatives of Mts protein for stimulating
 PT neuronal cell growth to treat neurological conditions e.g. Parkinson's
 PT disease, Alzheimer's disease, Down's syndrome and stroke.
 XX Disclosure; Page; 60pp; English.
 XX This invention relates to functional derivatives of an Mts protein,
 CC particularly Mts1-del75, and Mts1-4S. Mts 1 (also known as S100B4) is a
 CC member of the S100 gene family. Mts1 proteins and their functional
 CC derivatives are used to stimulate the growth of neuronal cells and in the
 CC treatment of neurological conditions including those characterised by
 CC neuronal degeneration, death or injury such as Parkinson's disease,
 CC Alzheimer's disease, Down's syndrome, stroke, cardiac arrest, sciatic
 CC crush, spinal cord injury, injury to sensory neurons or degenerative
 CC disease of the retina. The present sequence represents the murine Mts 1
 CC protein, from which the Mts proteins of the invention may be derived.
 CC NOTE: The present sequence is not shown in the specification but it has
 CC been derived from the sequence given in Genes Dev. 3, 1989 (referred to
 CC on page 10 of the specification)
 XX Sequence 101 AA;

Query Match 87.1%; Score 74; DB 4; Length 101;
 Best Local Similarity 86.7%; Pred. No. 0.00018;
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NEFFEGGPDKPRKK 15
 Db 87 NEFFEGGPDKPRKK 101
 RESULT 12
 ABP99368
 ID ABP99368 standard; protein; 758 AA.
 XX AC ABP99368;
 XX 24-MAR-2003 (first entry)
 XX Arabidopsis thaliana polypeptide SEQ ID NO 12.
 XX Arabidopsis thaliana; herbicidal; plant; growth regulator; transgenic;
 KW herbicide-resistant.
 XX Arabidopsis thaliana.
 OS WO200266660-A2.
 PN 29-AUG-2002.
 XX 13-FEB-2002; 2002WO-EP001466.
 XX 16-FEB-2001; 2001DE-01007843.
 PR 23-MAY-2001; 2001DE-01025537.
 XX (META-) METANOMICS GMBH & CO KGAA.
 XX Plesch G, Blau A, Daeschner K, Klein M;
 PI WPI; 2002-674953/72.
 DR N-PSDB; ABZ66819.
 XX Identifying herbicides and plant growth regulators, from ability to
 PT inhibit specific genes, also use of these genes to prepare herbicide-
 PT resistant transgenic plants.
 XX

PS Claim 1; Page 132-134; 224pp; German.
 XX The invention relates to identifying compounds (A) with herbicidal
 CC activity from their ability to reduce or block expression or activity of
 CC gene products of specific nucleic acid sequences (I) or amino acid
 CC sequences (II) encoded by (I). (A) and antagonists of proteins encoded by
 CC (I) are useful as herbicides and plant growth regulators. Overexpression
 CC of (I) in plants is used to produce herbicide-resistant plants and (I)
 CC may also be mutated to identify altered sequences that encode proteins
 CC resistant to herbicides. The present sequence is that of a polypeptide of
 CC the invention
 XX Sequence 758 AA;
 SQ Query Match 50.6%; Score 43; DB 5; Length 758;
 Best Local Similarity 70.0%; Pred. No. 1.9e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NEFFEGGPDK 10
 Db 119 SEFFGFPDE 128
 RESULT 13
 AAY12102
 ID AAY12102 standard; protein; 101 AA.
 XX AC AAY12102;
 XX 18-JUN-1999 (first entry)
 DT Human 5' EST secreted protein SEQ ID NO: 415.
 DE Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductive hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition.
 XX Homo sapiens.
 OS WO9906554-A2.
 PN 11-FEB-1999.
 XX 31-JUL-1998; 98WO-IB001238.
 PF 01-AUG-1997; 97US-00905134.
 PR (GEST) GENSET.
 PA Dumas Milne Edwards J, Duclert A, Lacroix B;
 PI WPI; 1999-153784/13.
 DR N-PSDB; AAX40935.
 DR New nucleic acids encoding human secreted proteins - obtained from cDNA
 PT libraries prepared from kidney, fetal kidney, dystrophic muscle, muscle
 PT and heart tissue.
 XX Claim 34; Page 521-522; 622pp; English.
 XX AAX40826 to AAX41093 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY01602 and
 CC AAY11994 to AAY12260, respectively. The proteins given represent the
 CC signal peptide and an N-terminal fragment of a secreted protein. The
 CC nucleic acid sequences can be used for producing secreted human gene
 CC products. They can also be used to develop products for diagnosis and
 CC therapy. The proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, reproductive hormone
 CC regulating activity, chemotactic/chemokinetic activity, haemostatic and

CC thrombolytic activity, receptor/ligand activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell
 XX Sequence 101 AA;
 SQ

Query Match 49.4%; Score 42; DB 2; Length 101;
 Best Local Similarity 57.1%; Pred. No. 36;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPKK 14
 ||||: |||:
 Db 36 NEFFQYTAQPKPK 49

RESULT 14
 ABB89552
 ID ABB89552 standard; protein; 141 AA.
 XX
 AC ABB89552;
 XX
 DT 24-MAY-2002 (first entry)
 XX
 DE Human polypeptide SEQ ID NO 1928.
 XX
 KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 XX antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 PN WO200190304-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-US016450.
 XX
 PR 19-MAY-2000; 2000US-0205515P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI; 2002-122018/16.
 DR N-PSDB; ABL89961.
 XX
 XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders.
 XX
 PS Claim 11; SEQ ID NO 1928; 2081pp + Sequence Listing; English.
 XX
 XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemia; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic

CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

SQ Sequence 141 AA;

Query Match 49.4%; Score 42; DB 5; Length 141;
 Best Local Similarity 57.1%; Pred. No. 50;
 Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPKK 14
 ||||: |||:
 Db 36 NEFFQYTAQPKPK 49

RESULT 15

AAG91772
 ID AAG91772 standard; protein; 759 AA.

XX

AC AAG91772;

XX

DT 26-SEP-2001 (first entry)

XX

DE C glutamicum protein fragment SEQ ID NO: 5526.

XX

KW Corynebacterium; amino acid synthesis; vitamin; saccharide;

KW organic acid synthesis.

XX

OS Corynebacterium glutamicum.

XX

PN EPI108790-A2.

XX

PD 20-JUN-2001.

XX

PF 18-DEC-2000; 2000EP-00127688.

XX

PR 16-DEC-1999; 99JP-00377484.

PR

07-APR-2000; 2000JP-00159162.

PR

03-AUG-2000; 2000JP-00280988.

XX

PA (KYOW) KYOWA HAKKO KOGYO KK.

XX

PI Nakagawa S, Mizoguchi H, Ando S, Hayashi M, Ochiai K, Yokoi H;

PI

Tateishi N, Senoh A, Ikeda M, Ozaki A;

XX

WPI; 2001-376931/40.

DR

N-PSDB; AAH66991.

XX

PT Novel polynucleotides derived from Corynebacterium bacteria, for identifying

PT

mutation point of a gene, measuring expression of a gene, analyzing

PT

expression profile or pattern of a gene and identifying homologous gene.

XX

PS Claim 17; SEQ ID NO 5526; 246pp + Sequence Listing; English.

XX

XX The present invention provides a number of nucleotide and protein

CC

sequences from the Corynebacterium bacterium Corynebacterium glutamicum. These

CC

are useful for identifying the mutation point of a gene derived from a

CC

mutant of corynebacterium bacterium, measuring expression amount and analysing

CC

the expression profile or expression pattern of a gene derived from

CC

Corynebacterium bacterium, and identifying a homologue of a gene derived from

CC

Corynebacterium bacterium. Corynebacterium bacteria are useful for producing amino

CC

acids, nucleic acids, vitamins, saccharides and organic acids,

CC

particularly L-lysine. The present sequence is a protein described in the

CC

exemplification of the invention. Note: The sequence data for this patent

CC

did not form part of the printed specification, but was obtained in

CC

electronic format directly from the European Patent Office

XX

Sequence 759 AA;

SQ

Query Match 49.4%; Score 42; DB 4; Length 759;

Best Local Similarity 63.6%; Pred. No. 2.7e+02;

Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 BFFEGFPDKQP 12
| | | | | : |
Db 462 EVFEGFPQRP 472

Search completed: September 9, 2004, 10:53:07
Job time : 32.5484 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:47:43 ; Search time 8.41935 Seconds
(without alignments)
91.977 Million cell updates/sec

Title: US-10-067-618-7

Perfect score: 85

Sequence: 1 NEFFEGFPDQPRKK 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/6CTUS_COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	100.0	15	1	US-08-190-560-7
2	85	100.0	15	1	US-08-469-277-7
3	85	100.0	15	2	US-08-468-946-7
4	85	100.0	15	2	US-08-468-942-7
5	85	100.0	15	4	US-09-298-625-7
6	85	100.0	101	1	US-08-190-560-2
7	85	100.0	101	1	US-08-469-277-2
8	85	100.0	101	2	US-08-468-946-2
9	85	100.0	101	4	US-08-468-942-2
10	85	100.0	101	4	US-09-298-625-2
11	42	49.4	460	4	US-09-540-236-3771
12	41	48.2	97	1	US-07-662-198B-2
13	41	48.2	318	4	US-09-489-847-328
14	41	48.2	379	4	US-09-402-532-1
15	41	48.2	379	4	US-09-402-532-3
16	41	48.2	382	4	US-09-402-532-4
17	41	48.2	411	4	US-09-402-532-2
18	41	48.2	412	4	US-09-402-532-6
19	41	48.2	412	4	US-09-402-532-8
20	41	48.2	414	4	US-09-402-532-5
21	41	48.2	427	4	US-09-591-095-24
22	41	48.2	444	4	US-09-402-532-7
23	40	47.1	398	4	US-09-252-991A-30286
24	40	47.1	554	4	US-09-252-991A-23813
25	39	45.9	144	4	US-09-134-001C-4218
26	39	45.9	237	4	US-08-970-133-5
27	39	45.9	237	4	US-09-294-545-5

28 39 45.9 451 3 US-08-969-644-8 Sequence 8, Appli
29 39 45.9 451 3 US-08-444-189-8 Sequence 8, Appli
30 39 45.9 451 3 US-08-468-544-8 Sequence 8, Appli
31 39 45.9 772 4 US-09-252-991A-30121 Sequence 30121, A
32 39 45.9 2227 3 US-08-475-886-2 Sequence 2, Appli
33 39 45.9 2227 3 US-08-475-886-4 Sequence 4, Appli
34 39 45.9 2227 3 US-08-475-886-6 Sequence 6, Appli
35 39 45.9 2227 3 US-08-397-232-2 Sequence 2, Appli
36 39 45.9 2227 3 US-08-397-232-4 Sequence 4, Appli
37 39 45.9 2227 3 US-09-171-387-2 Sequence 2, Appli
38 39 45.9 2227 4 US-09-653-493-2 Sequence 2, Appli
39 39 45.9 2227 4 US-09-653-493-4 Sequence 2, Appli
40 39 45.9 2227 4 US-09-653-499-6 Sequence 6, Appli
41 39 45.9 2227 4 US-10-104-966-12 Sequence 12, Appli
42 39 45.9 2227 4 US-10-135-988-2 Sequence 2, Appli
43 39 45.9 2227 4 US-10-135-988-4 Sequence 4, Appli
44 39 45.9 2227 4 US-10-135-988-6 Sequence 6, Appli
45 38 44.7 364 4 US-09-107-532A-5044 Sequence 5044, Ap

ALIGNMENTS

RESULT 1
US-08-190-560-7
; Sequence 7, Application US/08190560
; Patent No. 5798257
; GENERAL INFORMATION:
; APPLICANT: Zain, Saveeda
; APPLICANT: Lukenidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 301 SANS UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-190-560-7

Query Match 100.0%; Score 85; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDQPRKK 15
Db 1 NEFFEGFPDQPRKK 15

US-08-469-277-7
; Sequence 7, Application US/08469277
; Patent No. 5801142
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,946
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/190,560
; FILING DATE: 31-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-468-946-7
Query Match 100.0%; Score 85; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NEFFEGFPDKQPRKK 15
Db 1 NEFFEGFPDKQPRKK 15
RESULT 4
US-08-468-942-7
; Sequence 7, Application US/08468942
; Patent No. 5965360
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,942
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/190,560
; FILING DATE: 31-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-468-946-7

US-08-469-277-7
; Sequence 7, Application US/08469277
; Patent No. 5801142
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,277
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE:
; US-08-469-277-7
Query Match 100.0%; Score 85; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NEFFEGFPDKQPRKK 15
Db 1 NEFFEGFPDKQPRKK 15
RESULT 3
US-08-468-946-7
; Sequence 7, Application US/08468946
; Patent No. 5843686
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States

```
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-468-942-7

Query Match          100.0%; Score 85; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NEFEGFPDKQPRKK 15
Db 1 NEFEGFPDKQPRKK 15

RESULT 5
US-09-298-625-7
; Sequence 7, Application US/09298625
; Patent No. 6638504
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 7879ZYAII-Z
; CURRENT APPLICATION NUMBER: US/09/298,625
; CURRENT FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide
; US-09-298-625-7

Query Match          100.0%; Score 85; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 3.8e-08;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NEFEGFPDKQPRKK 15
Db 1 NEFEGFPDKQPRKK 15

RESULT 6
US-08-190-560-2
; Sequence 2, Application US/08190560
; Patent No. 5798257
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York

; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 15 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-468-942-7

Query Match          100.0%; Score 85; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NEFEGFPDKQPRKK 15
Db 87 NEFEGFPDKQPRKK 101

RESULT 7
US-08-469-277-2
; Sequence 2, Application US/08469277
; Patent No. 5801142
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,277
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 7879ZY
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 7879ZY
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
```

; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-277-2

Query Match 100.0%; Score 85; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
||| ||||| ||||| |||||
Db 87 NEFFEGFPDKQPRKK 101

RESULT 8

US-08-468-946-2
; Sequence 2, Application US/08468946
; Patent No. 5843686

; GENERAL INFORMATION:

; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,946
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/190,560
; FILING DATE: 31-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-946-2

Query Match 100.0%; Score 85; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
||| ||||| ||||| |||||
Db 87 NEFFEGFPDKQPRKK 101

Query Match 100.0%; Score 85; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
||| ||||| ||||| |||||
Db 87 NEFFEGFPDKQPRKK 101

RESULT 9

US-08-468-942-2
; Sequence 2, Application US/08468942
; Patent No. 5965360

; GENERAL INFORMATION:

; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,942
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/190,560
; FILING DATE: 31-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-942-2

Query Match

Best Local Similarity 100.0%; Score 85; DB 2; Length 101;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
||| ||||| ||||| |||||

Db 87 NEFFEGFPDKQPRKK 101

RESULT 10

US-09-298-625-2
; Sequence 2, Application US/09298625
; Patent No. 6638504

; GENERAL INFORMATION:

; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 78792YAI11-Z
; CURRENT APPLICATION NUMBER: US/09/298,625
; CURRENT FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600

Query Match 100.0%; Score 85; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
||| ||||| ||||| |||||
Db 87 NEFFEGFPDKQPRKK 101


```
; CURRENT APPLICATION NUMBER: US/09/402,532
; CURRENT FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/JP98/01643
; EARLIER FILING DATE: 1998-04-09
; EARLIER APPLICATION NUMBER: JP 10-010289
; EARLIER FILING DATE: 1998-01-22
; EARLIER APPLICATION NUMBER: JP 9-184885
; EARLIER FILING DATE: 1997-07-10
; EARLIER APPLICATION NUMBER: JP 9-093355
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE:
; SEQ ID NO 1
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Human
US-09-402-532-1

Query Match      48.2%; Score 41; DB 4; Length 379;
Best Local Similarity 45.5%; Pred. No. 34;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      3 FPEGFPDQPR 13
Db      310 YYESFPDRDPK 320

RESULT 15
US-09-402-532-3
; Sequence 3, Application US/09402532
; Patent No. 6498019
; GENERAL INFORMATION:
; APPLICANT: Taniyama, Yoshio
; TITLE OF INVENTION: NOVEL PROTEIN, ITS PRODUCTION AND USE
; FILE REFERENCE: 2456USOP
; CURRENT APPLICATION NUMBER: US/09/402,532
; CURRENT FILING DATE: 1999-10-04
; EARLIER APPLICATION NUMBER: PCT/JP98/01643
; EARLIER FILING DATE: 1998-04-09
; EARLIER APPLICATION NUMBER: JP 10-010289
; EARLIER FILING DATE: 1998-01-22
; EARLIER APPLICATION NUMBER: JP 9-184885
; EARLIER FILING DATE: 1997-07-10
; EARLIER APPLICATION NUMBER: JP 9-093355
; EARLIER FILING DATE: 1997-04-11
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE:
; SEQ ID NO 3
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Murine
US-09-402-532-3

Query Match      48.2%; Score 41; DB 4; Length 379;
Best Local Similarity 45.5%; Pred. No. 34;
Matches 5; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      3 FPEGFPDQPR 13
Db      310 YYESFPDRDPK 320

Search completed: September 9, 2004, 11:01:13
Job time : 9.41935 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:53:18 ; Search time 20.129 Seconds
(without alignments)
238.975 Million cell updates/sec

Title: US-10-067-618-7

Perfect score: 85

Sequence: 1 NEFFEGFPDQPRKK 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpa/PTC_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpa/PTCUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	15	13	US-10-067-618-7
2	85	100.0	15	13	US-10-135-152-7
3	85	100.0	101	9	US-09-393-433-1
4	85	100.0	101	9	US-09-781-509-1
5	85	100.0	101	12	US-10-087-192-1158
6	85	100.0	101	13	US-10-067-618-2
7	85	100.0	101	13	US-10-135-152-2
8	85	100.0	101	14	US-10-269-643-1
9	74	87.1	101	9	US-09-393-433-2
10	74	87.1	101	9	US-09-781-509-2
11	74	87.1	101	14	US-10-269-643-2
12	74	87.1	119	12	US-10-087-192-1155
13	49	57.6	116	12	US-10-424-599-199625
14	46	54.1	300	12	US-10-424-599-217218
15	42	49.4	175	15	US-10-264-237-1928

16 42 49.4 759 9 US-09-738-626-5526 Sequence 5526, Ap
17 42 49.4 1117 15 US-10-369-493-18965 Sequence 18965, A
18 42 49.4 1234 14 US-10-203-311A-6 Sequence 6, Appli
19 42 49.4 1268 12 US-10-276-774-2166 Sequence 2166, Ap
20 41 48.2 97 14 US-10-097-340-274 Sequence 274, App
21 41 48.2 97 14 US-10-171-311-206 Sequence 206, App
22 41 48.2 97 15 US-10-236-031B-40 Sequence 40, Appli
23 41 48.2 144 12 US-10-282-122A-46523 Sequence 46523, A
24 41 48.2 318 12 US-10-351-334-328 Sequence 328, App
25 41 48.2 379 14 US-10-323-051-1 Sequence 1, Appli
26 41 48.2 379 14 US-10-323-051-3 Sequence 3, Appli
27 41 48.2 382 14 US-10-323-051-4 Sequence 4, Appli
28 41 48.2 411 14 US-10-323-051-2 Sequence 2, Appli
29 41 48.2 412 9 US-09-978-295A-157 Sequence 157, App
30 41 48.2 412 9 US-09-978-697-157 Sequence 157, App
31 41 48.2 412 9 US-09-978-192A-157 Sequence 157, App
32 41 48.2 412 9 US-09-999-832A-157 Sequence 157, App
33 41 48.2 412 10 US-09-978-189-157 Sequence 157, App
34 41 48.2 412 10 US-09-978-608A-157 Sequence 157, App
35 41 48.2 412 10 US-09-978-585A-157 Sequence 157, App
36 41 48.2 412 10 US-09-978-191A-157 Sequence 157, App
37 41 48.2 412 10 US-09-978-403A-157 Sequence 157, App
38 41 48.2 412 10 US-09-978-564A-157 Sequence 157, App
39 41 48.2 412 10 US-09-999-833A-157 Sequence 157, App
40 41 48.2 412 10 US-09-981-915A-157 Sequence 157, App
41 41 48.2 412 10 US-09-978-824-157 Sequence 157, App
42 41 48.2 412 10 US-09-918-585A-157 Sequence 157, App
43 41 48.2 412 10 US-09-978-423A-157 Sequence 157, App
44 41 48.2 412 10 US-09-978-193A-157 Sequence 157, App
45 41 48.2 412 10 US-09-999-830A-157 Sequence 157, App

ALIGNMENTS

RESULT 1

US-10-067-618-7
; Sequence 7, Application US/10067618
; Publication No. US20020169110A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 7879ZYAII-Z
; CURRENT APPLICATION NUMBER: US/10/067,618
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/298,625
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin ver. 2.1
; SEQ ID NO 7
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide
US-10-067-618-7

Query Match 100.0%; Score 85; DB 13; Length 15;

Best Local Similarity 100.0%; Pred. No. 6.2e+07;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NEFFEGFPDQPRKK 15

Db 1 NEFFEGFPDQPRKK 15

```
RESULT 2
US-10-135-152-7
; Sequence 7, Application US/10135152
; Publication No. US20020172680A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 7879ZYAIIII-Z
; CURRENT APPLICATION NUMBER: US/10/135,152
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US/09/298,625
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide
US-10-135-152-7

Query Match          100.0%; Score 85; DB 13; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.2e-07; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
   |||||
Db 1 NEFFEGFPDKQPRKK 15

RESULT 3
US-09-393-433-1
; Sequence 1, Application US/09393433
; Patent No. US20010011126A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/09/393,433
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-393-433-1

Query Match          100.0%; Score 85; DB 9; Length 101;
Best Local Similarity 100.0%; Pred. No. 4.6e-06; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
   |||||
Db 87 NEFFEGFPDKQPRKK 101

RESULT 4
US-09-781-509-1
; Sequence 1, Application US/09781509
; Patent No. US20020099010A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/09/781,509
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 09/393,433
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-781-509-1

Query Match          100.0%; Score 85; DB 9; Length 101;
Best Local Similarity 100.0%; Pred. No. 4.6e-06; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
   |||||
Db 87 NEFFEGFPDKQPRKK 101

RESULT 5
US-10-087-192-1158
; Sequence 1158, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1158
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-1158

Query Match          100.0%; Score 85; DB 12; Length 101;
Best Local Similarity 100.0%; Pred. No. 4.6e-06; Indels 0; Gaps 0;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
   |||||
Db 87 NEFFEGFPDKQPRKK 101

RESULT 6
US-10-067-618-2
; Sequence 2, Application US/10067618
; Publication No. US20020169110A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 7879ZYAIIII-Z
; CURRENT APPLICATION NUMBER: US/10/067,618
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/298,625
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 08/468,942
```

```
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-067-618-2

Query Match      100.0%; Score 85; DB 13; Length 101;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
   |||||
Db 87 NEFFEGFPDKQPRKK 101

RESULT 7
US-10-135-152-2
; Sequence 2, Application US/10135152
; Publication No. US20020172680A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 78792A111-2
; CURRENT APPLICATION NUMBER: US/10/135,152
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US/09/298,625
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-135-152-2

Query Match      100.0%; Score 85; DB 13; Length 101;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
   |||||
Db 87 NEFFEGFPDKQPRKK 101

RESULT 8
US-10-269-643-1
; Sequence 1, Application US/10269643
; Publication No. US20030100503A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/10/269,643
; CURRENT FILING DATE: 2002-10-11
```

```
; PRIOR APPLICATION NUMBER: US/09/781,509
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 09/393,433
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-269-643-1

Query Match      100.0%; Score 85; DB 14; Length 101;
Best Local Similarity 100.0%; Pred. No. 4.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
   |||||
Db 87 NEFFEGFPDKQPRKK 101

RESULT 9
US-09-393-433-2
; Sequence 2, Application US/09393433
; Patent No. US20010011126A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/09/393,433
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-393-433-2

Query Match      87.1%; Score 74; DB 9; Length 101;
Best Local Similarity 86.7%; Pred. No. 0.00028;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
   |||||
Db 87 NEFFEGFPDKQPRKK 101

RESULT 10
US-09-781-509-2
; Sequence 2, Application US/09781509
; Patent No. US20020099010A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/09/781,509
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 09/393,433
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-781-509-2

Query Match      87.1%; Score 74; DB 9; Length 101;
```

```
Best Local Similarity 86.7%; Pred. No. 0.00028;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRK 15
   ||||| |||:||||
Db 87 NEFFEGCPDKPRK 101

RESULT 11
US-10-269-643-2
; Sequence 2, Application US/10269643
; Publication No. US20030100503A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berzen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/10/269,643
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/781,509
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 09/393,433
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-269-643-2

Query Match 87.1%; Score 74; DB 14; Length 101;
Best Local Similarity 86.7%; Pred. No. 0.00028;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRK 15
   ||||| |||:||||
Db 87 NEFFEGCPDKPRK 101

RESULT 12
US-10-087-192-1155
; Sequence 1155, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-1155

Query Match 87.1%; Score 74; DB 12; Length 119;
Best Local Similarity 86.7%; Pred. No. 0.00033;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRK 15
   ||||| |||:||||
Db 105 NEFFEGCPDKPRK 119

RESULT 13
US-10-424-599-199625
; Sequence 199625, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 199625
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_22286C.1.pep
US-10-424-599-199625

Query Match 57.6%; Score 49; DB 12; Length 116;
Best Local Similarity 64.3%; Pred. No. 3.5;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRK 14
   ||||| |||:||||
Db 91 NDAFEGKSPDKPRK 104

RESULT 14
US-10-424-599-217218
; Sequence 217218, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 217218
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_38175C.1.pep
US-10-424-599-217218

Query Match 54.1%; Score 46; DB 12; Length 300;
Best Local Similarity 57.1%; Pred. No. 29;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRK 14
   ||||| |||:||||
Db 275 NDAFEGKSPDKPRK 288

RESULT 15
US-10-264-237-1928
; Sequence 1928, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
```

```
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 1928
; LENGTH: 175
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (63)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (144)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-1928
```

```
Query Match          49.4%; Score 42; DB 15; Length 175;
Best Local Similarity 57.1%; Pred. No. 73;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
```

```
Qy      1 NEFFEGFDDKQPRK 14
Db      70 NEFFQYTAPKQPRK 83
```

```
Search completed: September 9, 2004, 11:04:47
Job time : 21.129 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:44:32 ; Search time 7.64516 Seconds
(without alignments)
188.730 Million cell updates/sec

Title: US-10-067-618-7

Perfect score: 85

Sequence: 1 NEFFEGFDKQPRKK 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_78:*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	85	100.0	100	2 A53217	placental calcium-
2	85	100.0	101	2 A48219	calvasculin - huma
3	74	87.1	15	2 I49407	placental calcium-
4	74	87.1	101	2 S06207	calvasculin - mous
5	74	87.1	101	2 S01759	calvasculin - rat
6	47	55.3	282	2 A85076	hypothetical prote
7	45	52.9	376	2 S64113	ARC1 protein - yea
8	44	51.8	370	2 C95374	probable ABC trans
9	43	50.6	457	2 T09290	late embryonic abu
10	42	49.4	1234	2 T00363	hypothetical prote
11	42	49.4	1708	2 A81866	WD-40 repeat prote
12	41	48.2	98	2 A41988	S-100 calcium-bind
13	41	48.2	244	2 H83780	NADPH-flavin oxido
14	41	48.2	272	2 T4755	hypothetical prote
15	41	48.2	304	2 T48281	hypothetical prote
16	41	48.2	504	2 G02474	interferon regulat
17	41	48.2	504	2 T13302	probable primase -
18	41	48.2	749	2 T50397	probable serine/th
19	41	48.2	1484	2 C97196	probable membrane
20	40	47.1	161	2 S68771	bacterioferritin -
21	40	47.1	161	2 AG3597	bacterioferritin (
22	40	47.1	166	2 H64605	hypothetical prote
23	40	47.1	206	2 G71218	hypothetical prote
24	40	47.1	217	2 JX0095	glutathione transf
25	40	47.1	260	2 T98503	crbf protein - Ent
26	40	47.1	276	2 B71907	hypothetical prote
27	40	47.1	436	2 G97701	polynucleotide ade
28	40	47.1	448	2 E86291	hypothetical prote
29	40	47.1	501	2 T49474	hypothetical prote

30 40 47.1 575 2 F71400 hypothetical prote
31 40 47.1 1506 2 T32909 hypothetical prote
32 39 45.9 169 2 A84320 hypothetical prote
33 39 45.9 221 2 T36174 hypothetical prote
34 39 45.9 237 2 S64718 formin-binding pro
35 39 45.9 306 2 T11042 hypothetical prote
36 39 45.9 398 2 S06324 dnab protein homol
37 39 45.9 406 2 G97708 hypothetical prote
38 39 45.9 451 2 A37386 dnab protein homol
39 39 45.9 493 2 T01495 hypothetical prote
40 39 45.9 620 2 T50232 actin-like protein
41 39 45.9 746 2 S31816 genome polyprotein
42 39 45.9 762 2 H83415 cis/trans isomerase
43 39 45.9 864 1 S63229 probable multifunc
44 39 45.9 926 2 T04679 hypothetical prote
45 39 45.9 1004 2 A48821 Wnt-5 protein - fr

ALIGNMENTS

RESULT 1

A53217

Placental calcium-binding protein homolog - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 17-Mar-1999

C:Accession: A53217

R:Polans, A.S.; Palczewski, K.; Asson-Batres, M.A.; Ohguro, H.; Witkowska, D.; Haley, T.
J. Biol. Chem. 269, 6233-6240, 1994

A:Title: Purification and primary structure of capl, an S-100-related calcium-binding pr

A:Reference number: A53217; MUID:94164991; PMID:8119967

A:Accession: A53217

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-100 <POL>

C:Superfamily: S-100 protein; calmodulin repeat homology

C:Keywords: acetylated amino end; calcium binding; EF hand

F:6-40/domain: calmodulin repeat homology <EFL>

F:1/Modified site: acetylated amino end (Ala) #status experimental

Query Match

Best Local Similarity 100.0%; Score 85; DB 2; Length 100;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFDKQPRKK 15

DB 86 NEFFEGFDKQPRKK 100

RESULT 2

A48219

calvasculin - human

N:Alternate names: calcium-binding protein L (CAPL); metastasin; placental calcium-bindin

C:Species: Homo sapiens (man)

C>Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 13-Aug-1999

C:Accession: A48219; B44470; F01223

R:Engelkamp, D.; Schaefer, B.W.; Mattei, M.G.; Erne, P.; Heizmann, C.W.

Proc. Natl. Acad. Sci. U.S.A. 90, 6547-6551, 1993

A:Title: Six S100 genes are clustered on human chromosome 1q21: identification of two ge

A:Reference number: A48219; MUID:93342029; PMID:8341667

A:Accession: A48219

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 1-101 <ENG>

A:Cross-references: GB:218950; NID:G396708; PIDN:CAA79474.1; PID:G396710

R:Engelkamp, D.; Schaefer, B.W.; Erne, P.; Heizmann, C.W.

Biochemistry 31, 10258-10264, 1992

A:Title: S100 alpha, CAPL, and CACY: molecular cloning and expression analysis of three (

A:Reference number: A44470; MUID:93041710; PMID:11384693

A:Accession: B44470

A>Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-101 <EN2>

A;Cross-references: GB:M80563; GB:M77499; NID:g179916; PIDN:AAA51920.1; PID:g179917
 A;Experimental source: heart
 A;Note: sequence extracted from NCBI backbone (NCBIP:116496)
 R;Tomida, Y.; Terasawa, M.; Kobayashi, R.; Hidaka, H.
 Biochem. Biophys. Res. Commun. 189, 1310-1316, 1992
 A;Title: Calcyclin and calvasculin exist in human platelets.
 A;Reference number: PC1222; MUID:93129189; PMID:1482346
 A;Accession: PC1223
 A;Molecule type: protein
 A;Residues: 8-18;36-57 <TOM>
 A;Experimental source: platelets
 C;Genetics:
 A;Gene: GDB:S100A4; CAPL
 A;Cross-references: GDB:119748; OMIM:114210
 A;Map position: 1q21-1q21
 A;Introns: 47/3
 C;Superfamily: S-100 protein; calmodulin repeat homology
 C;Keywords: calcium binding; EF hand
 F;7-41/Domain: calmodulin repeat homology <EF1>
 F;50-82/Domain: calmodulin repeat homology <EF2>
 Query Match 100.0%; Score 85; DB 2; Length 101;
 Best Local Similarity 100.0%; Pred. No. 7.6e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 NEFFEGFPDKQPRKK 15
 |||||
 DB 87 NEFFEGFPDKQPRKK 101
 |||||
 RESULT 3
 149407
 placental calcium-binding protein - western wild mouse (fragment)
 C;Species: Mus spretus (western wild mouse)
 C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 29-Sep-1999
 C;Accession: I49407
 R;Ko, M.S.; Wang, X.; Horton, J.H.; Hagen, M.D.; Takahashi, N.; Maezaki, Y.; Nadeau, J.H.
 Mamm. Genome 5, 349-355, 1994
 A;Title: Genetic mapping of 40 cDNA clones on the mouse genome by PCR.
 A;Reference number: I48934; MUID:94319082; PMID:8043949
 A;Accession: I49407
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-15 <RES>
 A;Cross-references: EMBL:U05696; NID:g497016; PIDN:AAA61936.1; PID:g497017
 C;Superfamily: S-100 protein; calmodulin repeat homology
 C;Keywords: calcium binding; EF hand
 Query Match 87.1%; Score 74; DB 2; Length 15;
 Best Local Similarity 86.7%; Pred. No. 6.6e-06;
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NEFFEGFPDKQPRKK 15
 |||||
 DB 1 NEFFEGCPDKPRKK 15
 |||||
 RESULT 4
 S06207
 calvasculin - mouse
 N;Alternate names: calcium-binding protein mts1; calcium-binding protein pEL98; placenta
 C;Species: Mus musculus (house mouse)
 C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 20-Jun-2000
 C;Accession: S06207; JH0097; S07981; A26803; A41411; I48674
 R;Ebraldize, A.; Tulchinsky, E.; Grigorian, M.; Afanasyeva, A.; Senin, V.; Ravazova, E.
 Genes Dev. 3, 1086-1093, 1989
 A;Title: Isolation and characterization of a gene specifically expressed in different me
 A;Reference number: S06207; MUID:89378739; PMID:2550322
 A;Accession: S06207
 A;Molecule type: mRNA
 A;Residues: 1-101 <EBR>
 A;Cross-references: EMBL:X16190; NID:g54926; PIDN:CAA34316.1; PID:g54927

R;Tulchinsky, E.M.; Grigorian, M.S.; Ebraldize, A.K.; Milshina, N.I.; Lukanidin, E.M.
 Gene 87, 219-223, 1990
 A;Title: Structure of gene mts1, transcribed in metastatic mouse tumor cells.
 A;Reference number: JH0097; MUID:90236313; PMID:2332170
 A;Accession: JH0097
 A;Molecule type: DNA
 A;Residues: 1-101 <TUL>
 A;Cross-references: GB:M36578; GB:M36579
 A;Experimental source: liver
 R;Tulchinsky, B.
 submitted to the EMBL Data Library, August 1989
 A;Reference number: S07981
 A;Accession: S07981
 A;Molecule type: DNA
 A;Residues: 1-47; 'VSGSXFNQ', 56-57, 'RTDEAA' <TU2>
 A;Cross-references: EMBL:X16094; NID:g53249; PIDN:CAA34224.1; PID:g53250
 R;Jackson-Grusby, L.L.; Swiergiel, J.; Linzer, D.I.H.
 Nucleic Acids Res. 15, 6677-6690, 1987
 A;Title: A growth-related mRNA in cultured mouse cells encodes a placental calcium bindi
 A;Reference number: A26803; MUID:87316927; PMID:3628004
 A;Accession: A26803
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-101 <JAC>
 A;Cross-references: GB:X05835; NID:g50310; PIDN:CAA29282.1; PID:g50311
 R;Soto, K.; Endo, H.; Fujiyoshi, T.
 J. Biochem. 103, 48-53, 1988
 A;Title: Cloning of the sequences expressed abundantly in established cell lines: identi
 A;Reference number: A41411; MUID:88198109; PMID:3162911
 A;Accession: A41411
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-101 <GOT>
 A;Cross-references: GB:D00208; NID:g220569; PIDN:BAA00148.1; PID:g220570
 R;Tulchinsky, E.; Kzamerov, D.; Ford, H.L.; Reshetnyak, E.; Lukanidin, E.; Zain, S.
 Oncogene 8, 79-86, 1993
 A;Title: Characterization of a positive regulatory element in the mts1 gene.
 A;Reference number: I48674; MUID:93141279; PMID:8423998
 A;Accession: I48674
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-47; 'VSGSXFNQ', 48-54 <RES>
 A;Cross-references: EMBL:X16094; NID:g53249; PIDN:CAA34224.1; PID:g53250
 C;Comment: Gene mts1 is expressed in metastatic cells.
 C;Genetics:
 A;Gene: mts1
 A;Introns: 47/3
 C;Superfamily: S-100 protein; calmodulin repeat homology
 C;Keywords: calcium binding; cancer; EF hand
 F;7-41/Domain: calmodulin repeat homology <EF1>
 F;50-82/Domain: calmodulin repeat homology <EF2>
 Query Match 87.1%; Score 74; DB 2; Length 101;
 Best Local Similarity 86.7%; Pred. No. 4.9e-05;
 Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 NEFFEGFPDKQPRKK 15
 |||||
 DB 87 NEFFEGCPDKPRKK 101
 |||||
 RESULT 5
 S01759
 calvasculin - rat
 N;Alternate names: calcium-binding protein p9Ka/42A; gene p9Ka protein
 C;Species: Rattus norvegicus (Norway rat)
 C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 13-Aug-1999
 C;Accession: S01759; B31373; S26496; S26497
 R;Barracough, R.; Savin, J.; Dube, S.K.; Rudland, P.S.
 J. Mol. Biol. 198, 13-20, 1987
 A;Title: Molecular cloning and sequence of the gene for p9Ka : a cultured myoepithelial
 A;Reference number: S01759; MUID:88118907; PMID:3430604
 A;Accession: S01759

A:Molecule type: DNA
A:Residues: 1-101 <BAR>
A:Cross-references: EMBL:X06916; NID:G56832; PIDN:CAA30014.1; PID:G56833
R:Maslakowski, P.; Shooter, E.M.
Proc. Natl. Acad. Sci. U.S.A. 85, 1277-1281, 1988
A:Title: Nerve growth factor induces the genes for two proteins related to a family of cation channels
A:Reference number: A94189; MUID:88125019; PMID:3422491
A:Accession: B31373
A:Molecule type: mRNA
A:Residues: 1-101 <MAS>
A:Cross-references: EMBL:X03628; NID:G206829; PIDN:AAA42098.1; PID:G206830
R:de Voege, M.W.; Mukherjee, B.B.
Oncogene 7, 109-119, 1992
A:Title: Transformation of normal rat kidney cells by v-K-ras enhances expression of transforming growth factor- α
A:Reference number: S26496; MUID:192158347; PMID:1741158
A:Accession: S26496
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 65-100 <DEV>
A:Cross-references: EMBL:X64023
A:Accession: S26497
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 3-97 <DE2>
A:Cross-references: EMBL:X64022
C:Genetics:
A:Gene: p9Ka
A:Introns: 47/3
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; EF hand
F:50-82/Domain: calmodulin repeat homology <EF2>
Query Match 87.1%; Score 74; DB 2; Length 101;
Best Local Similarity 86.7%; Pred. No. 4.9e-05;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NEFFEGFPDKPRKK 15
|||||:|||||
Db 87 NEFFEGCPDKPRKK 101
|||||:|||||
RESULT 6
A85076
Hypothetical protein AT4g07710 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: A85076
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Harbor Laboratory
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488; PMID:10617198
A:Accession: A85076
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-282 <STO>
A:Cross-references: GB:NC_001268; NID:G7267369; PIDN:CAB77930.1; GSPDB:GN00140
C:Genetics:
A:Gene: At4g07710
A:Map position: 4
Query Match 55.3%; Score 47; DB 2; Length 282;
Best Local Similarity 50.0%; Pred. No. 3.8;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 NEFFEGFPDKPRKK 14
|||||:|||||
Db 17 NRYEGFPDQTHR 30
|||||:|||||
RESULT 7
S64113
ARCL1 protein - yeast (Saccharomyces cerevisiae)
N/Alternate names: G4Pl protein; protein G3085; protein YGL105W

C:Species: Saccharomyces cerevisiae
C:Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 09-Dec-2002
A:Accession: S64113; S69425; S72274
R:Castagnoli, L.; Paoluzi, S.; Minenkova, O.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64112
A:Accession: S64113
A:Molecule type: DNA
A:Residues: 1-376 <CAS>
A:Cross-references: EMBL:Z72627; NID:G1322647; PIDN:CAA96812.1; PID:G1322648; MIPS:YGL10719
A:Experimental source: strain S288C
R:Castagnoli, L.; Paoluzi, S.; Minenkova, O.
submitted to the EMBL Data Library, April 1996
A:Reference number: S69417
A:Accession: S69425
A:Molecule type: DNA
A:Residues: 1-376 <CAW>
A:Cross-references: EMBL:X97644; NID:G1310710; PIDN:CAA66247.1; PID:G1310719
R:Simos, G.; Segref, A.; Pasolo, F.; Hellmuth, K.; Shevchenko, A.; Mann, M.; Hurt, E.C.
EMBO J. 15, 5437-5448, 1996
A:Title: The yeast protein Arc1p binds to tRNA and functions as a cofactor for the methyltransferase Arc1p
A:Reference number: S72274; MUID:97050848; PMID:8895587
A:Accession: S72274
A>Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-376 <SIM>
A:Cross-references: EMBL:X95481; NID:G1620459; PIDN:CAA64750.1; PID:G1620460
C:Genetics:
A:Gene: SGD:ARCL1; G4Pl
A:Cross-references: MIPS:YGL105W; SGD:S0003073
A:Map position: 7L
A:Complex: homodimer
C:Function:
A:Description: functions as a cofactor for methionyl- and glutamyl-tRNA synthetases
C:Superfamily: endothelial monocyte-activating polypeptide II precursor (pro-EMAP II)
C:Keywords: cytosol; homodimer
Query Match 52.9%; Score 45; DB 2; Length 376;
Best Local Similarity 61.5%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 3 FFEFGFPDKPRKK 15
|||||:|||||
Db 307 FFEFGDEAPMKQ 319
|||||:|||||
RESULT 8
C95374
Probable ABC transporter, ATP-binding protein SMA1646 [imported] - Sinorhizobium meliloti
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 17-May-2002
C:Accession: C95374
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: C95374
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-370 <KUR>
A:Cross-references: GB:AE006469; PIDN:AAK65557.1; PID:G14524035; GSPDB:GN00165
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Fuhrer, A.; Abola, P.; Ampe, F.; Barloy-Hubler, F.;
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:

A;Gene: SMal646
A;Genome: plasmid
C;Superfamily: inner membrane protein malk; ATP-binding cassette homology

Query Match 51.8%; Score 44; DB 2; Length 370;
Best Local Similarity 63.6%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 FEGFPDKQPRK 14
|||:::|
Db 17 FRGLNRPQPK 27

RESULT 9
T09290
late embryonic abundant protein EMB8 - white spruce
C;Species: Picea glauca (white spruce)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 12-Jun-2003
C;Accession: T09290
R;Dong, J.Z.; Dunstan, D.I.
submitted to the EMBL Data Library, June 1996
A;Description: Cloning and characterization of 6 novel plant embryogenesis-associated genes
A;Reference number: Z16629
A;Accession: T09290
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-457 <DON>
A;Cross-references: EMBL:L47118; NID:g1350544; PID:g1350545
C;Genetics:
A;Gene: EMB8
C;Superfamily: alpha/beta hydrolase

Query Match 50.6%; Score 43; DB 2; Length 457;
Best Local Similarity 58.3%; Pred. No. 28;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 NEFFEGPDKQ 12
:::|:::|
Db 30 SQFRAGPDKKP 41

RESULT 10
T00363
hypothetical protein KIAA0674 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 21-Jul-2000
C;Accession: T00363
R;Ishikawa, K.; Nagase, T.; Suyama, M.; Miyajima, N.; Tanaka, A.; Kotani, H.; Nomura, N.
DNA Res. 5, 169-176, 1998
A;Title: Prediction of the coding sequences of unidentified human genes. X. The complete
A;Reference number: Z14142; MUID:98403980; PMID:9734811
A;Accession: T00363
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1234 <ISH>
A;Cross-references: EMBL:AB014574; NID:G3327161; PIDN:BA031649.1; PID:G3327162
A;Experimental source: brain; clone HK02519
C;Genetics:
A;Gene: KIAA0674
C;Superfamily: BKBP-type peptidylprolyl isomerase homology
F;212-263/Domain: BKBP-type peptidylprolyl isomerase homology <PPI>

Query Match 49.4%; Score 42; DB 2; Length 1234;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 NEFFEGPDKQPRK 14
|||:::|
Db 51 NEFFQYAPKQPK 64

RESULT 11
AE1866

WD-40 repeat protein [imported] - Nostoc sp. (strain PCC 7120)
C;Species: Nostoc sp. PCC 7120
A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C;Accession: AE1866
R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguch
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
A;Reference number: AB1807; MUID:21595285; PMID:11759840
A;Accession: AE1866
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1708 <KUR>
A;Cross-references: GB:BA000019; PIDN:BA072436.1; PID:g17129823; GSPDB:GN00179
A;Experimental source: strain PCC 7120
C;Genetics:
A;Gene: all0478

Query Match 49.4%; Score 42; DB 2; Length 1708;
Best Local Similarity 50.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 EFFEFGPDKQPRK 15
|||:::|
Db 681 EFFEFGYSDROQKKE 694

RESULT 12
A41988
S-100 calcium-binding protein A2 - human
N;Alternate names: calcium-binding protein Can19; S-100 calcium-binding protein L
C;Species: Homo sapiens (man)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 29-Aug-1997
C;Accession: A41988
R;Lee, S.W.; Tomasetto, C.; Swisshelm, K.; Keyomarsi, K.; Sager, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 2504-2508, 1992
A;Title: Down-regulation of a member of the S100 gene family in mammary carcinoma cells
A;Reference number: A41988; MUID:92196147; PMID:1372446
A;Accession: A41988
A;Molecule type: mRNA
A;Residues: 1-98 <LEE>
A;Cross-references: GB:M87068; NID:g179896
A;Experimental source: mammary carcinoma cells
A;Note: sequence extracted from NCBI backbone (NCBIN:88561, NCBIIP:88562)
C;Genetics:
A;Gene: GDB:S100A2; CAN19; S100L
A;Cross-references: GDB:211118; OMIM:176993
A;Map position: 1q21-1q21
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: calcium binding; EF hand
F;8-42/Domain: calmodulin repeat homology <EF1>
F;51-83/Domain: calmodulin repeat homology <EF2>

Query Match 48.2%; Score 41; DB 2; Length 98;
Best Local Similarity 60.0%; Pred. No. 12;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 NEFFEGPDK 10
|||:::|
Db 88 NDFFQGCPCR 97

RESULT 13
H83780
NADPH-flavin oxidoreductase BH1048 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C;Accession: H83780
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132

Search completed: September 9, 2004, 10:59:42
Job time : 8.64516 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:32:57 ; Search time 4.25806 Seconds
(without alignments)
183.429 Million cell updates/sec

Title: US-10-067-618-7

Perfect score: 85

Sequence: 1 NEFFEGFPDQPRKK 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	100	1 S104_BOVIN	P35466 bos taurus
2	85	100.0	101	1 S104_HUMAN	P26447 homo sapien
3	74	87.1	101	1 S104_MOUSE	P07091 mus musculus
4	74	87.1	101	1 S104_RAT	P05942 rattus norv
5	45	52.9	376	1 GAP1_YEAST	P46672 saccharomyc
6	43	50.6	457	1 EMB8_PICGL	Q40863 picea glauc
7	41	48.2	97	1 S102_HUMAN	P29034 homo sapien
8	41	48.2	242	1 UR1E_CHLTE	Q8kf69 chlorobium
9	41	48.2	498	1 IRF5_HUMAN	Q13568 homo sapien
10	40	47.1	161	1 BFR_BRUME	P49944 brucella me
11	40	47.1	206	1 Y010_PYRHO	O57770 pyrococcus
12	40	47.1	217	1 GTMU_CAVPO	P16413 cavia porce
13	40	47.1	757	1 HT16_HYDAT	P53356 hydra atten
14	39	45.9	620	1 ARP8_SCHPO	Q9us07 schizosacch
15	39	45.9	745	1 CHSD_ASPFU	P78746 aspergillus
16	39	45.9	864	1 FAS_YEAST	P53848 s foliaci
17	39	45.9	1010	1 WNT5_DROME	P28466 drosophila
18	39	45.9	2226	1 POLG_HPAV2	P26580 hepatitis a
19	39	45.9	2226	1 POLG_HPAV4	P26581 hepatitis a
20	39	45.9	2226	1 POLG_HPAV8	P26582 hepatitis a
21	39	45.9	2227	1 POLG_HPAVH	P08617 hepatitis a
22	39	45.9	2227	1 POLG_HPAVL	P08441 hepatitis a
23	39	45.9	2227	1 POLG_HPAVM	P13901 hepatitis a
24	39	45.9	2230	1 POLG_HPAVS	P14553 simian hepa
25	38.5	45.3	521	1 GM12_METAC	Q8tiy2 methanosarc
26	38	44.7	145	1 Y01W_BACSU	P54534 bacillus su
27	38	44.7	225	1 GSTZ_EUPES	P57108 euphorbia e
28	38	44.7	254	1 K2CZ_SCHPO	Q94281 schizosacch
29	38	44.7	342	1 ME18_MOUSE	P23798 mus musculu
30	38	44.7	537	1 APE3_YEAST	P37302 saccharomyc
31	38	44.7	626	1 PEPO_LACLA	Q07744 lactococcus
32	38	44.7	626	1 PEPO_LACLC	Q09145 lactococcus
33	38	44.7	980	1 PEX6_HUMAN	Q13608 homo sapien

RESULT 1

ID	S104_BOVIN	STANDARD;	PRT;	100 AA.
AC	P35466;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Placental calcium-binding protein homolog.			
GN	S100A4 OR CAPL.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Retina;			
RX	MEDLINE=94164991; PubMed=8119967;			
RA	Polans A.S., Palczewski K., Asson-Batres M.A., Ohguro H., Witowska D.,			
RA	Haley T.L., Baizer L., Crabb J.W.;			
RT	"Purification and primary structure of Capl, an S-100-related			
RT	calcium-binding protein isolated from bovine retina.";			
RL	J. Biol. Chem. 269:6233-6240(1994).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98430466; PubMed=9759666;			
RA	Duarte W.R., Kasugai S., Iimura T., Oida S., Takenaga K., Ohya K.,			
RA	Ishikawa I.;			
RT	"cDNA cloning of S100 calcium-binding proteins from bovine			
RT	periodontal ligament and their expression in oral tissues.";			
RL	J. Dent. Res. 77:1694-1699(1998).			
CC	!- SIMILARITY: Belongs to the S-100 family.			
CC	!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; D89056; BAA13754.1; ..			
DR	PIR; A53217; A53217.			
DR	HSSP; P30801; 1A03.			
DR	InterPro; IPR001751; CAPP_S100.			
DR	InterPro; IPR002048; EF-hand.			
DR	Pfam; PF00036; ehand; 1.			
DR	Pfam; PF01023; S100; 1.			
DR	ProDom; PD003407; CAPP_S100; 1.			
DR	ProDom; PD000012; EF-hand; 1.			
DR	PROSITE; PS00018; EF HAND; 1.			
DR	PROSITE; PS00303; S100_CAPP; 1.			
KW	Calcium-binding; Acetylation.			
FT	INIT MET 0			
FT	MOD_RES 1 1 ACETYLATION.			

ALIGNMENTS

```

FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SQ SEQUENCE 100 AA; 11675 MW; DFFCAA7561D5EFA4 CRC64;

Query Match 100.0%; Score 85; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
Db 86 NEFFEGFPDKQPRKK 100

RESULT 2
ID S104_HUMAN STANDARD; PRT; 101 AA.
AC P26447;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Placental calcium-binding protein (Calvasculin) (S100 calcium-binding
DE Protein A4) (MTS1 protein).
GN S100A4 OR CAPL OR MTS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=93041710; PubMed=1384693;
RA Engelkamp D., Schaefer B.W., Erne P., Heizmann C.W.;
RT "S100 alpha, CAPL, and CACY: molecular cloning and expression
RT analysis of three calcium-binding proteins from human heart.";
RL Biochemistry 31:10258-10264(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93342029; PubMed=8341667;
RA Engelkamp D., Schaefer B., Mattei M.-G., Erne P., Heizmann C.W.;
RT "Six S100 genes are clustered on human chromosome 1q21:
RT identification of two genes coding for the two previously unreported
RT calcium-binding proteins S100D and S100E.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:6547-6551(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=93028421; PubMed=1329089;
RA Tulchinsky E.M., Ford H.L., Kramerov D., Reshetnyak E., Grigorian M.,
RA Zain S., Lukanidin E.;
RT "Transcriptional analysis of the mts1 gene with specific reference to
RT 5' flanking sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:9146-9150(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix, and Prostate;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

```

```

RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M80563; AAA51920.1; -
CC EMBL; Z18950; CAA79474.1; -
CC EMBL; Z33457; CAA83880.1; -
CC EMBL; BC000838; AAH00838.1; -
CC EMBL; BC016300; AAH16300.1; -
CC FIR; A48219; A48219.
CC PDB; 1MJ1; 30-OCT-02.
CC SWISS-2DPAGE; P26447; HUMAN.
CC Genew; HGNC:10494; S100A4.
CC MIM; 114210; -
CC GO; GO:0005509; F:calcium ion binding; TAS.
CC InterPro; IPR001751; CaBP_S100.
CC InterPro; IPR002048; EF-hand.
CC Pfam; PF00036; ehand; 1.
CC ProDom; PD003407; CaBP_S100; 1.
CC ProDom; PD000012; EF-hand; 1.
CC PROSITE; PS00018; EF-hand; 1.
CC PROSITE; PS00303; S100_CaBP; 1.
KW Calcium-binding; 3D-structure.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SQ SEQUENCE 101 AA; 11728 MW; 286D2B7B07EDB562 CRC64;

Query Match 100.0%; Score 85; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRKK 15
Db 87 NEFFEGFPDKQPRKK 101

RESULT 3
ID S104_MOUSE STANDARD; PRT; 101 AA.
AC P07091; P20066;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Placental calcium-binding protein (18A2) (PEL98) (MTS1 protein)
DE (Metastatic cell protein).
GN S100A4 OR CAPL OR MTS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;
RX MEDLINE=87316927; PubMed=3628004;
RA Jackson-Grusby L.H., Swiergiel J., Linzer D.I.H.;
RT "A growth-related mRNA in cultured mouse cells encodes a placental
RT calcium binding protein.";
RL Nucleic Acids Res. 15:6677-6690(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/C;

```


RA MEDLINE=88198109; PubMed=3162911;
RX Goto K., Endo H., Fujiyoshi T.;
RT "Cloning of the sequences expressed abundantly in established cell
RT lines: identification of a cDNA clone highly homologous to S-100, a
RT calcium binding protein.";
RL J. Biochem. 103:48-53 (1988).
RN [3]
RN SEQUENCE FROM N.A.
RP MEDLINE=89378739; PubMed=2550322;
RX Ravalidze A., Tulchinsky E., Grigorian M., Afanasyeva A., Senin V.,
RA Revazova E., Lukanidin E.;
RT "Isolation and characterization of a gene specifically expressed in
RT different metastatic cells and whose deduced gene product has a high
RT degree of homology to a Ca2+-binding protein family.";
RL Genes Dev. 3:1086-1093 (1989).
RN [4]
RN SEQUENCE FROM N.A.
RP SEQUENCE=Liver;
RC TISSUE=Liver;
RX MEDLINE=90236313; PubMed=2332170;
RA Tulchinsky E.M., Grigorian M.S., Ebralidze A.K., Milshina N.I.,
RA Lukanidin E.M.;
RT "Structure of gene mts1, transcribed in metastatic mouse tumor
RT cells.";
RL Gene 87:219-223 (1990).
RN [5]
RN SEQUENCE OF 1-54 FROM N.A.
RP MEDLINE=93141279; PubMed=8423998;
RX Tulchinsky E., Kramerov D., Ford H.L., Reshetnyak E., Lukanidin E.,
RA Zain S.;
RT "Characterization of a positive regulatory element in the mts1 gene.";
RL Oncogene 8:79-86 (1993).
CC -!- TISSUE SPECIFICITY: Specifically expressed in different metastatic
CC cells.
CC -!- INDUCTION: The mRNA coding for this protein increases in
CC abundance after serum stimulation of quiescent mouse fibroblasts.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X05835; CAA29282.1; -;
CC EMBL; D00208; BAA00148.1; -;
CC EMBL; M36578; AAA39749.1; -;
CC EMBL; M36579; AAA39750.1; -;
CC EMBL; X16190; CAA34316.1; -;
CC EMBL; X16094; CAA34224.1; -;
CC PIR; S06207; S06207.
CC HSP; P30801; 1A03.
CC MGD; MGI:1330282; S100a4.
CC InterPro; IPR001751; CaBP_S100.
CC InterPro; IPR002048; EF-hand.
CC Pfam; PF00036; ehand; 1.
CC Pfam; PF01023; S_100; 1.
CC ProDom; PD003407; CaBP_S100; 1.
CC PROSITE; PS00018; EF_HAND; 1.
CC PROSITE; PS00303; S100_CABP; 1.
CC Calcium-binding; Pfam.
CC CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
CC CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
CC CONFLICT 47 47 G -> GVSGSXFNQ (IN REF. 5).
CC SEQUENCE 101 AA; 11721 MW; 2302254B674C873 CRC64;
Query Match 87.1%; Score 74; DB 1; Length 101;
Best Local Similarity 86.7%; Pred. No. 1e-05;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NEFFEGFPDKPRKK 15

Db 87 NEFFEGFPDKPRKK 101
||||| |||:||||
S104 RAT
ID S104 RAT STANDARD; PRT; 101 AA.
AC P05942;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Placental calcium-binding protein (Nerve growth factor induced protein
DE 42A) (P9K).
DE S100A4.
GN Rattus norvegicus (Rat).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=88125019; PubMed=3422491;
RA Maslakowski P., Shooter E.M.;
RT "Nerve growth factor induces the genes for two proteins related to a
RT family of calcium-binding proteins in PC12 cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:1277-1281 (1988).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=88118907; PubMed=3430604;
RA Barracough R., Savin J., Dube S.K., Rudland P.S.;
RT "Molecular cloning and sequence of the gene for p9Ka. A cultured
RT myoepithelial cell protein with strong homology to S-100, a calcium-
RT binding protein.";
RL J. Mol. Biol. 198:13-20 (1987).
RN [3]
RN SEQUENCE OF 3-101 FROM N.A.
RX MEDLINE=92158347; PubMed=1741158;
RA de Vonge M.W., Mukherjee B.B.;
RT "Transformation of normal rat kidney cells by v-K-ras enhances
RT expression of transin 2 and an S-100-related calcium-binding
RT protein.";
RL Oncogene 7:109-119 (1992).
CC -!- INDUCTION: By nerve growth factor.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X06916; CAA30014.1; -;
CC EMBL; J03628; AAA42098.1; -;
CC EMBL; X64022; -; NOT_ANNOTATED_CDS.
CC EMBL; X64023; -; NOT_ANNOTATED_CDS.
CC PIR; S01759; S01759.
CC HSP; P30801; 1A03.
CC InterPro; IPR001751; CaBP_S100.
CC InterPro; IPR002048; EF-hand.
CC Pfam; PF00036; ehand; 1.
CC Pfam; PF01023; S_100; 1.
CC ProDom; PD003407; CaBP_S100; 1.
CC PROSITE; PS00018; EF_HAND; 1.
CC PROSITE; PS00303; S100_CABP; 1.
CC Calcium-binding.
CC CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
CC CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
CC CONFLICT 47 47 G -> GVSGSXFNQ (IN REF. 5).
CC SEQUENCE 101 AA; 11776 MW; EA0619CEB4F487C1 CRC64;
Query Match 87.1%; Score 74; DB 1; Length 101;
Best Local Similarity 86.7%; Pred. No. 1e-05;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 NEFFEGFPDKPRKK 15

DR	InterPro; IPR008232; EMAP11.
DR	InterPro; IPR008994; Nucleic acid_OB.
DR	InterPro; IPR002547; tRNA_bind.
DR	Pfam; PF01588; tRNA_bind_1.
DR	PIRSF; PIRSF005381; EMAP11; 1.
DR	PROSITE; PS00886; TRBD; 1.
KW	DNA-binding; RNA-binding; tRNA-binding.
FT	DOMAIN 205 307
FT	CONFLICT 137 149
FT	CONFLICT 181 182
FT	SEQUENCE 376 AA; 42084 MW; BOCDEAD564900C7 CRC64;
FT	REF. 1).
FT	KK->LL (IN REF. 1).
FT	TRNA-BINDING.
FT	PAGGAADAATAAKAD -> LRVALMLQQGS (IN
FT	REF. 1).
FT	Score 45; DB 1; Length 376;
FT	Best Local Similarity 61.5%; Pred. No. 3.6;
FT	Mismatches 2; Indels 0; Gaps 0;
FT	Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY	3 FFEFGPDKQPKK 15
DB	: :
DB	307 FFEFGDPAFMKQ 319
RESULT 6	
EMB8_PICGL	
ID_EMB8_PICGL	STANDARD; PRT; 457 AA.
AC Q40863;	
DT 01-NOV-1997 (Rel. 35, Created)	
DT 01-NOV-1997 (Rel. 35, Last sequence update)	
DT 28-FEB-2003 (Rel. 41, Last annotation update)	
DE Embryogenesis-associated protein EMB8.	
GN EMB8	
OS Picea glauca (White spruce).	
OC Eukaryota; Viridiplantae; Streptophyta;	
OC Spermatophyta; Coniferopsida; Pinaceae; Picea.	
NCBI_TaxID=3330;	
[1]	
SEQUENCE FROM N.A., AND DEVELOPMENTAL STAGE.	
RX MEDLINE=99277601; PubMed=10350098;	
RA Dong J.-Z., Dunstan D.I.;	
RT "Cloning and characterization of six embryogenesis-associated cDNAs	
RT from somatic embryos of Picea glauca and their comparative expression	
RT during zygotic embryogenesis.";	
RL Plant Mol. Biol. 39:859-864(1999).	
CC -I- DEVELOPMENTAL STAGE: In somatic embryos, expressed at low levels	
CC during the first phase of maturation and is most abundant 14 days	
CC after maturation begins. Not detected in zygotic embryos.	
CC -I- SIMILARITY: Belongs to the UPF0017 family.	
CC -----	
CC This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC the European Bioinformatics Institute. There are no restrictions on its	
CC use by non-profit institutions as long as its content is in no way	
CC modified and this statement is not removed. Usage by and for commercial	
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC or send an email to licenses@isb-sib.ch).	
CC -----	
CC EMBL; L47118; AAB01571.1; --	
CC PIR; T09290; T09290.	
DR InterPro; IPR000379; Ser_estrs.	
DR InterPro; IPR000952; UPF0017.	
DR PROSITE; PS01133; UPF0017; 1.	
SQ SEQUENCE 457 AA; 51019 MW; 5E6483F80F582A7A CRC64;	
Query Match	50.6%; Score 43; DB 1; Length 457;
Best Local Similarity	58.3%; Pred. No. 9.9;
Matches	7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
QY 1 NEFFEGFDKPQ 12	
DB :: :	
DB 30 SQFRAGFDKPP 41	
RESULT 7	

S102 HUMAN
ID S102 HUMAN STANDARD; PRT; 97 AA.
AC P29034; 000266; Q9BU83;
DT 01-DEC-1992 (Rel. 24, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE S100 calcium-binding protein A2 (S-100L protein) (CAN19).
GN S100A2 OR S100L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92196147; PubMed=1372446;
RZ Lee S.W., Tomasetto C., Swisshelm K., Keyomarsi K., Sager R.;
RT "Down-regulation of a member of the S100 gene family in mammary
RT carcinoma cells and reexpression by azadeoxycytidine treatment.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2504-2508(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98142474; PubMed=9481475;
RZ Wicki R., Franz C., Scholl F.A., Heizmann C.W., Schaefer B.W.;
RT "Repression of the candidate tumor suppressor gene S100A2 in breast
RT cancer is mediated by site-specific hypermethylation.";
RL Cell Calcium 22:243-254(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary;
RX MEDLINE=22388257; PubMed=12477932;
RZ Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McWeary P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE OF 36-39 AND 41-48.
RC TISSUE=Keratinocytes;
RX MEDLINE=93162043; PubMed=1286667;
RZ Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
RA Vandekerckhove J.;
RT "Microsequences of 145 proteins recorded in the two-dimensional gel
RT protein database of normal human epidermal keratinocytes.";
RL Electrophoresis 13:960-969(1992).
CC -!- FUNCTION: May act as a modulator against excess calcium
CC accumulation in normal human mammary epithelial cells. May also
CC play a role in suppressing tumor cell growth.
CC -!- TISSUE SPECIFICITY: A subset of epithelial cells amongst which
CC normal human mammary epithelial cells and keratinocytes.
CC -!- DEVELOPMENTAL STAGE: Preferentially expressed in normal human
CC mammary epithelial cells as opposed to tumor-derived ones. The
CC level of S100L was shown to correlate inversely with tumor
CC progression.
CC -!- INDUCTION: By growth factors in early G1 phase and probably by
CC cell-cycle regulation in S phase. DNA methylation probably plays
CC a direct negative role in suppressing S100L gene expression in
CC tumor cells.
CC -!- MISCELLANEOUS: This protein binds two calcium ions (By

CC similarity). Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; M87068; -; NOT ANNOTATED_CDS.
DR EMBL; Y07755; CAA69033.1; -;
DR EMBL; BC002829; AAH02829.2; -;
DR HSPG; P30801; 1A03.
DR Aarhus/Ghent-2DPAGE; 9027; IEF.
DR Aarhus/Ghent-2DPAGE; 9051; IEF.
DR Genew; HGNC:10492; S100A2.
DR MIM; 176993; -;
DR GO; GO:0005509; F:calcium ion binding; NAS.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CBP; 1.
KW Calcium-binding.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT CONFLICT 61 61 S->N (IN REF.1).
SQ SEQUENCE 97 AA; 10986 MW; CFE06CD8DC08D250 CRC64;
Query Match 48.2%; Score 41; DB 1; Length 97;
Best Local Similarity 60.0%; Pred. No. 3.9;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
Qy 1 NEFFEGPDK 10
Db 87 NDFFGQCPDR 96
RESULT 8
UBIE CHLTE STANDARD; PRT; 242 AA.
AC Q8KF69;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Menquinone biosynthesis methyltransferase ubiB (EC 2.1.1.-).
GN UBIE OR CT0462.
OS Chlorobium tepidum.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=TLS / ATCC 49652 / DSM 12025;
RZ MEDLINE=22103685; PubMed=12093901;
RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Fang F.,
RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA Nierman W.C., Feldblum T.V., Hansen C.L., Craven M.B., Radune D.,
RA Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT "The complete genome sequence of Chlorobium tepidum TLS, a
RT photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
CC -!- FUNCTION: Methyltransferase required for the conversion of
CC dimethylmenaquinone (DMKH2) to menaquinone (MKH2) (By similarity).
CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =

```

CC      s-adenosyl-L-homocysteine + menaquinol.
CC      -!- PATHWAY: Menaquinone biosynthesis; last step.
CC      -!- SIMILARITY: Belongs to the ubie family.
CC
CC      This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch)
CC
CC      EMBL; AE012823; AM71705.1; --
CC      TIGR; CT0462; --; 1.
CC      HAMAP; MF 01813; --; 1.
CC      InterPro; IPR000051; SAM bind.
CC      InterPro; IPR004033; Ubie/COQ5 Metrf.
CC      Pfam; PF01209; Ubie_methyltran; 1.
CC      PROSITE; PS01183; Ubie_1; FALSE NEG.
CC      PROSITE; PS01184; Ubie_2; FALSE NEG.
CC      Menaquinone biosynthesis; Transferase; Methyltransferase;
KW      Complete proteome.
SQ      SEQUENCE 242 AA; 27307 MW; A516BC8D4DC4135A CRC64;

Query Match      48.2%; Score 41; DB 1; Length 242;
Best Local Similarity 63.6%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY      2 EFEGFPDKQP 12
      ||| ||| |||
DB      110 EFLEGAEKVP 120

RESULT 9
IRFS_HUMAN
ID      IRF5_HUMAN STANDARD; PRT; 498 AA.
AC      Q13568; Q9BQF0;
DT      01-NOV-1997 (Rel. 35, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Interferon regulatory factor 5 (IRF-5).
GN      IRF5.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Grossman A., Mitrucker H.W., Lantonio L., Mak T.W.;
RL      Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RA      Tissue=Kidney;
RX      MEDLINE=22386257; PubMed=12477932;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Ditschenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA      Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
RA      Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
RA      Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length
RT      human and mouse cDNA sequences."

```

```

RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC      -!- SUBCELLULAR LOCATION: Nuclear.
CC      -!- SIMILARITY: Belongs to the IRF family.
CC
CC      This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch)
CC
CC      EMBL; U51127; AAA96056.1; --
CC      EMBL; BC004201; AAH04201.1; --
CC      EMBL; BC004139; AAH04139.1; --
CC      PIR; G02474; G02474.
CC      HSSP; P23906; 1IRP.
CC      TRANSFAC; T05104; --
CC      Genew; HGNC:6120; IRF5.
CC      MIM; 607218; --
CC      InterPro; IPR001346; IRF.
CC      InterPro; IPR008984; SMAD_FHA.
CC      Pfam; PF06055; IRF; 1.
CC      PRINTS; PR00267; INTPRNEGFC.
CC      ProDom; PD002355; IRF; 1.
CC      SMART; SM00348; IRF; 1.
CC      PROSITE; PS00601; IRF; 1.
KW      Transcription regulation; DNA-binding; Nuclear protein.
FT      DNA_BIND 16 118 TRYPTOPHAN PENTAD REPEAT.
FT      DOMAIN 142 149 POLY-GLU.
FT      CONFLICT 161 175 EDVKWPPTLQPPTLR -> DAQSGPHMTPTYSLLKEDVKW
      (IN REF. 1).
SQ      SEQUENCE 498 AA; 56044 MW; 01B2ED95C28384E8 CRC64;

Query Match      48.2%; Score 41; DB 1; Length 498;
Best Local Similarity 53.8%; Pred. No. 24;
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY      3 RFEGFPDKQPKK 15
      ||| ||| ||| |||
DB      399 FGEEWPDKPKREK 411

RESULT 10
BFR_BRUME
ID      BFR_BRUME STANDARD; PRT; 161 AA.
AC      P49944;
DT      01-OCT-1996 (Rel. 34, Created)
DT      01-OCT-1996 (Rel. 34, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Bacterioferritin (BFR).
GN      BFR OR BMEII0704.
OS      Brucella melitensis.
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC      Brucellaceae; Brucella.
OX      NCBI_TaxID=29459;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      STRAIN=16M / ATCC 23456 / Biotype 1;
RX      MEDLINE=95212556; PubMed=7698330;
RA      Denoel P.A., Zygmunt M.S., Weynants V.E., Tibor A., Lichtfouse B.,
RA      Briffeuil P., Limet J.N., Letesson J.J.;
RT      "Cloning and sequencing of the bacterioferritin gene of Brucella
RT      melitensis 16M strain."
RL      FEBS Lett. 361:238-242 (1995).
RN      [2]
RP      SEQUENCE FROM N.A.
RA      STRAIN=16M / ATCC 23456 / Biotype 1;
RX      MEDLINE=20020109; PubMed=11756588;
RA      Delvecchio V.G., Kapatal V., Redkar R.J., Patra G., Mujar C., Los T.,
RA      Ivanova N., Anderson I., Bhattacharyya A., Lykidis A., Reznik G.,
RA      Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltsman E.,
RA      Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,

```

```
RA Haselkorn R., Kyrpides N., Overbeek R.;
RT "the genome sequence of the facultative intracellular pathogen
RL Brucella melitensis.";
CC -1- FUNCTION: May perform analogous functions in iron detoxification
CC and storage to that of animal ferritins.
CC -1- COFACTOR: Binds 1 heme B (iron-protoporphyrin IX) group per dimer
CC (By similarity).
CC -1- SUBUNIT: Oligomer of 24 identical subunits (By similarity).
CC -1- MISCELLANEOUS: The di-iron binding site functions as active site
CC where iron ions are oxidized from iron(II) to iron(III) before
CC they are stored (By similarity).
CC -1- SIMILARITY: Belongs to the bacterioferritin family.
CC -1- SIMILARITY: Contains 1 ferritin-like diiron domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U19760; AAC43322.1; -.
CC PIR; AG3597; AG3597.
CC PIR; S68771; S68771.
CC HSP; P11056; IBCF.
CC InterPro; IPR002024; Bacterioferritin.
CC InterPro; IPR008331; Ferritin Dps.
CC InterPro; IPR009040; Ferritin_like.
CC Pfam; PF00210; ferritin; 1.
CC PRINTS; PR00601; BACFERRITIN.
CC ProDom; PD002269; Bacterioferritin; 1.
CC TIGRFAMs; TIGR00754; bfr; 1.
CC PROSITE; PS00549; BACTERIOFERRITIN; 1.
CC PROSITE; PS0905; FERRITIN LIKE; 1.
CC Iron storage; Heme; Metal-binding; Iron; Complete proteome.
FT DOMAIN 1 145 FERRITIN-LIKE DIIRON.
FT METAL 18 18 IRON 1 (BY SIMILARITY).
FT METAL 51 51 IRON 1 (BY SIMILARITY).
FT METAL 51 51 IRON 2 (BY SIMILARITY).
FT METAL 52 52 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT METAL 54 54 IRON 1 (BY SIMILARITY).
FT METAL 94 94 IRON 2 (BY SIMILARITY).
FT METAL 127 127 IRON 1 (BY SIMILARITY).
FT METAL 127 127 IRON 2 (BY SIMILARITY).
FT METAL 130 130 IRON 2 (BY SIMILARITY).
SQ SEQUENCE 161 AA; 18659 MW; 5A92C1759B68BD1B CRC64;

Query Match 47.1%; Score 40; DB 1; Length 161;
Best Local Similarity 77.8%; Pred. No. 10;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 FFEFGPDQK 11
Db 64 FFEFGFNLQ 72

RESULT 11
Y010 PYRHO
ID Y010 PYRHO STANDARD; PRT; 206 AA.
AC OS7770;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein PH0010.
GN PH0010.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.F., Horikoshi K.,
RA Masuchi Y., Shizuwa H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76 (1998).
CC -1- SIMILARITY: Contains 1 AMMECR1 domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP000001; BAA29078.1; -.
CC PIR; G71218; G71218.
CC HAMAP; MF_00645; -.
CC InterPro; IPR002733; DUF51.
CC Pfam; PF01871; AMMECR1; 1.
CC ProDom; PD009671; DUF51; 1.
CC TIGRFAMs; TIGR00296; TIGR00296; 1.
CC Hypothetical protein; Complete proteome.
SQ SEQUENCE 206 AA; 23623 MW; 5338BBF798F770D1 CRC64;

Query Match 47.1%; Score 40; DB 1; Length 206;
Best Local Similarity 53.8%; Pred. No. 13;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 EFEGPDQKPRK 14
Db 115 ELLEGPEERPRK 127

RESULT 12
GTMU CAVPO
ID GTMU CAVPO STANDARD; PRT; 217 AA.
AC P16413;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Glutathione S-transferase B (EC 2.5.1.18) (GST B) (GST class-mu).
GN GSTM1.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE
RC TISSUE=Liver;
RX MEDLINE=90236961; PubMed=2332413;
RA Kamei K., Oshino R., Hara S.;
RT "Amino acid sequence of glutathione S-transferase b from guinea pig
RT liver.";
RL J. Biochem. 107:111-117(1990).
CC -1- FUNCTION: Conjugation of reduced glutathione to a wide number of
CC exogenous and endogenous hydrophobic electrophiles.
CC -1- CATALYTIC ACTIVITY: RX + glutathione = HX + R-S-glutathione.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: Belongs to the GST superfamily. Mu family.
CC PIR; JX0095; JX0095.
DR HSP; P04905; 2GST.
DR InterPro; IPR004046; GST_Cterm.
DR InterPro; IPR003081; GST_mu.
DR InterPro; IPR004045; GST_Nterm.
```

```

DR Pfam; PF00043; GST_C; 1.
DR Pfam; PF02798; GST_N; 1.
DR PRINTS; PR01267; GSTRNSPASEM.
KW Transferase; Multigene family.
SQ SEQUENCE 217 AA; 25719 MW; D29F7951D4E9365E CRC64;

Query Match          47.1%; Score 40; DB 1; Length 217;
Best Local Similarity 77.8%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      2 EFEGFDPK 10
      || || || ||
Db      125 EFLEGIPDK 133

RESULT 13
ID HT16 HYDAT STANDARD; PRT; 757 AA.
AC P53356;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase HTK16 (EC 2.7.1.112).
GN HTK16.
OS Hydra attenuata (Hydra) (Hydra vulgaris).
OC Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Anthomedusae;
OC Hydridae; Hydra.
OX NCBI_TaxID=6087;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Irvine.
RX MEDLINE=94181280; PubMed=8134129;
RA Chan T.A., Chu C.A., Rauen K.A., Krother M., Tatarszewicz S.M.,
RA Steele R.E.;
RT "Identification of a gene encoding a novel protein-tyrosine kinase
RT containing SH2 domains and ankyrin-like repeats.";
RL Oncogene 9,1253-1259(1994).
CC -!- FUNCTION: May be involved in signal transduction.
CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -!- TISSUE SPECIFICITY: Epithelial cells.
CC -!- SIMILARITY: Belongs to the Tyr family of protein kinases.
CC -!- SIMILARITY: Contains 2 SH2 domains.
CC -!- SIMILARITY: Contains 5 ANK repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U00936; AAC27350.1; -.
CC HSSP; P08631; 1AD5.
CC InterPro; IPR002110; ANK.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR000980; SH2.
CC InterPro; IPR001245; Tyr_kinase.
CC InterPro; IPR008266; Tyr_kinase_AS.
CC Pfam; PF00023; ank; 5.
CC Pfam; PF00069; pkinase; 1.
CC Pfam; PF00017; SH2; 2.
CC PRINTS; PR00401; SH2DOMAIN.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC ProDom; PD000093; SH2; 2.
CC SMART; SM00248; ANK; 5.
CC SMART; SM00252; SH2; 2.
CC SMART; SM00219; TyrKc; 1.
CC PROSITE; PS50088; ANK_REPEAT; 2.
CC PROSITE; PS0297; ANK_REPEAT_REGION; 1.
CC PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.

```

```

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE; PS50001; SH2; 2.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation;
KW SH2 domain; ANK repeat; Repeat.
FT DOMAIN 10 102 SH2 1.
FT REPEAT 115 147 ANK 1.
FT REPEAT 151 180 ANK 2.
FT REPEAT 184 214 ANK 3.
FT REPEAT 219 248 ANK 4.
FT REPEAT 252 281 ANK 5.
FT DOMAIN 287 379 SH2 2.
FT DOMAIN 484 740 PROTEIN_KINASE.
FT NP_BIND 490 498 ATP (BY SIMILARITY).
FT BINDING 516 516 ATP (BY SIMILARITY).
FT ACT_SITE 608 608 BY SIMILARITY.
FT MOD_RES 746 746 PHOSPHORYLATION (POTENTIAL).
SQ SEQUENCE 757 AA; 85598 MW; 144D09E9109D794F CRC64;

```

```

Query Match          47.1%; Score 40; DB 1; Length 757;
Best Local Similarity 50.0%; Pred. No. 55;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```
QY      2 EFEGFDPKQPK 13
      |||: :|||:
Db      270 EFFDYPVNPQPK 281

```

```

RESULT 14
ARP8 SCHPO STANDARD; PRT; 620 AA.
ID ARP8 SCHPO STANDARD; PRT; 620 AA.
AC Q9US07; Q9U033;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Probable actin-related protein 8.
GN SPAC664.02C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomyces.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odeh C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Leleau V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Porashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002).
RN [2]

```

```

RP SEQUENCE OF 377-590 FROM N.A., AND SUBCELLULAR LOCATION.
RC STRAIN=968 h90;
RA MEDLINE=20223868; PubMed=10759889;
RX Ding D.-Q., Tomita Y., Yamamoto A., Chikashige Y., Haraguchi T.,
RT Hiraoka Y.;
RT "Large-scale screening of intracellular protein localization in living
RT fission yeast cells by the use of a GFP-fusion genomic DNA library.";
RL Genes Cells 5:169-190(2000).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: Belongs to the actin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AL136235; CAB65803.1; -.
CC EMBL; AB027842; BAA87146.1; -.
CC PIR; T50232; T50232.
CC GeneDB SPombe; SPAC664.02c; -.
CC InterPro; IPR004000; Actin_like.
CC Pfam; PF00022; actin; 1.
CC SMART; SM00268; ACTIN; 1.
CC Structural protein; Cytoskeleton; Nuclear protein.
KW CONFLICT 377 381 NSFFS -> IASPR (IN REF. 2).
FT SEQUENCE 620 AA; 70057 MW; B5E071464189244F CRC64;
SQ
Query Match 45.9%; Score 39; DB 1; Length 620;
Best Local Similarity 57.1%; Pred. No. 66;
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPRK 14
Db ||| ||| |||
377 NSFFSRTPDKGTRK 390

RESULT 15
CHSD ASPFU STANDARD; PRT; 745 AA.
AC P78746;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Chitin synthase D (EC 2.4.1.16) (Chitin-UDP acetyl-glucosaminyl
DE transferase D) (Class-VI chitin synthase D).
GN CHSD.
OS Aspergillus fumigatus (Sartorya fumigata).
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5085;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=237;
RX MEDLINE=96401428; PubMed=8807804;
RA Mellado E., Specht C.A., Robbins P.W., Holden D.W.;
RT "Cloning and characterization of chsD, a chitin synthase-like gene of
RT Aspergillus fumigatus.";
RL FEMS Microbiol. Lett. 143:69-76(1996).
CC -!- FUNCTION: Plays a major role in cell wall biogenesis.
CC -!- CATALYTIC ACTIVITY: UDP-N-acetyl-D-glucosamine + {(1,4)-(N-acetyl-
CC beta-D-glucosaminyl)}(N) = UDP + {(1,4)-(N-acetyl-beta-D-
CC glucosaminyl)}(N+1).
CC -!- SUBCELLULAR LOCATION: Plasma membrane-bound.
CC -!- SIMILARITY: BELONGS TO THE CHITIN SYNTHASE FAMILY. SUBFAMILY CLASS
CC VI.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U62614; AAB60781.1; -.
CC InterPro; IPR004835; Chitin synthase.
CC InterPro; IPR001173; GlycoTrans_2.
CC Pfam; PF03142; Chitin synthase_2; 1.
KW Transferase; Glycosyltransferase; Transmembrane; Cell wall;
KW Multigene family.
FT TRANSMEM 26 46 POTENTIAL.
FT TRANSMEM 55 75 POTENTIAL.
FT TRANSMEM 412 432 POTENTIAL.
FT TRANSMEM 434 454 POTENTIAL.
FT TRANSMEM 464 484 POTENTIAL.
SQ SEQUENCE 745 AA; 84521 MW; B43B5D2B2D4A9632 CRC64;
Query Match 45.9%; Score 39; DB 1; Length 745;
Best Local Similarity 46.2%; Pred. No. 80;
Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 NEFFEGFPDKQPR 13
Db ||: : |||
616 NDNMKYPVDEQPR 628

Search completed: September 9, 2004, 10:53:52
Job time : 5.25806 secs

```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:42:57 ; Search time 24.9677 Seconds
(without alignments)
189.556 Million cell updates/sec

Title: US-10-067-618-7
Perfect score: 85
Sequence: 1 NEFFEGPDKPRKK 15

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85	100.0	101	6 Q9TV56	Q9TV56 canis famil
2	74	87.1	15	11 Q62531	Q62531 mus spretus
3	47	55.3	282	10 Q9M0T7	Q9M0T7 arabidopsis
4	45	52.9	425	5 Q8I560	Q8I560 plasmodium
5	44	51.8	170	9 Q8W6G7	Q8W6G7 sinorhizobi
6	44	51.8	262	17 Q8TPF9	Q8TPF9 methanosarc
7	44	51.8	370	16 Q92Y11	Q92Y11 rhizobium m
8	43	50.6	282	16 Q8EWA0	Q8EWA0 mycoplasma
9	43	50.6	330	11 Q9CQ51	Q9CQ51 mus musculus
10	43	50.6	378	17 Q8Q0I4	Q8Q0I4 methanosarc
11	43	50.6	758	10 Q9FN73	Q9FN73 arabidopsis
12	43	50.6	758	10 Q8S9K3	Q8S9K3 arabidopsis
13	43	50.6	997	12 Q8U5Z0	Q8U5Z0 simian herp
14	43	50.6	997	12 Q7T5D7	Q7T5D7 simian herp
15	42	49.4	625	16 Q98BN5	Q98BN5 rhizobium l
16	42	49.4	644	11 Q80YW7	Q80YW7 mus musculus

17	42	49.4	650	11 Q80YW6	Q80YW6 mus musculus
18	42	49.4	739	11 Q80YW8	Q80YW8 mus musculus
19	42	49.4	756	11 Q80YW9	Q80YW9 mus musculus
20	42	49.4	759	16 Q8NPG9	Q8NPG9 corynebacte
21	42	49.4	1234	4 Q9Y4D0	Q9Y4D0 homo sapien
22	42	49.4	1708	16 Q8Y2I2	Q8Y2I2 anabaena sp
23	41	48.2	96	16 Q87GF3	Q87GF3 vibrio para
24	41	48.2	120	17 Q8PS60	Q8PS60 methanosarc
25	41	48.2	137	16 Q818U0	Q818U0 bacillus ce
26	41	48.2	144	16 Q81M74	Q81M74 bacillus an
27	41	48.2	192	4 Q9NPQ6	Q9NPQ6 homo sapien
28	41	48.2	229	2 Q9RPN0	Q9RPN0 versinia pe
29	41	48.2	243	16 Q8A7R7	Q8A7R7 bacteroides
30	41	48.2	244	16 Q9KE10	Q9KE10 bacillus ha
31	41	48.2	272	4 Q9UG04	Q9UG04 homo sapien
32	41	48.2	304	10 Q9LZ42	Q9LZ42 arabidopsis
33	41	48.2	319	10 Q8LDS4	Q8LDS4 arabidopsis
34	41	48.2	405	2 Q8GBR5	Q8GBR5 treponema m
35	41	48.2	412	4 Q8NCC3	Q8NCC3 homo sapien
36	41	48.2	412	4 Q9Y2B3	Q9Y2B3 homo sapien
37	41	48.2	412	11 Q8VEB4	Q8VEB4 mus musculus
38	41	48.2	504	9 Q9MBX2	Q9MBX2 streptococc
39	41	48.2	506	9 Q9ZXG7	Q9ZXG7 streptococc
40	41	48.2	525	16 Q82HM9	Q82HM9 streptomyce
41	41	48.2	568	16 Q8F3E2	Q8F3E2 leptospira
42	41	48.2	644	10 Q8RY45	Q8RY45 arabidopsis
43	41	48.2	648	10 Q8S5G6	Q8S5G6 oryza sativ
44	41	48.2	655	10 Q9FID4	Q9FID4 arabidopsis
45	41	48.2	749	3 Q9P7X5	Q9P7X5 schizosacch

ALIGNMENTS

RESULT 1

Q9TV56 PRELIMINARY; PRT; 101 AA.

AC Q9TV56; 01-MAY-2000 (TREMREL. 13, Created)

DT 01-MAY-2000 (TREMREL. 13, Last sequence update)

DT 01-OCT-2003 (TREMREL. 25, Last annotation update)

DE Metastasin.

GN Mtsl.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI_TaxID=9615;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Madin-Darby; TISSUE=Kidney;

RA Miyamori H., Hasegawa K., Kim K., Sato H.;

RT "Expression of metastasin associated mtsl gene is co-induced with membrane type-1 matrix metalloproteinase (MT1-MMP) during oncogenic transformation and tubular formation of madin darby canine kidney (MDCK) epithelial cells."

RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.

DR EMBL; AB031064; BAA83419.1; -.

DR HSP; P30801; 1A03.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR InterPro; IPR001753; CaBP_S100.

DR InterPro; IPR002048; EF-hand.

DR Pfam; PF00036; ehand; 1.

DR Pfam; PF01023; S_100; 1.

DR ProDom; PD003407; CaBP_S100; 1.

DR ProDom; PD000012; EF-hand; 1.

DR PROSITE; PS000018; EF_HAND; 1.

DR PROSITE; PS00303; S100_CaBP; 1.

SQ SEQUENCE 101 AA; 11847 MW; 5535387DB7577DF0 CRC64;

Query Match 100.0%; Score 85; DB 6; Length 101;
Best Local Similarity 100.0%; Pred. No. 5.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Thu Sep 9 12:03:27 2004

```
QY 1 NEFFEGFDPKQPRK 15
   |||||
Db 87 NEFFEGFDPKQPRK 101

RESULT 2
Q62531 PRELIMINARY; PRT; 15 AA.
AC Q62531;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Placental calcium-binding protein (18A2) (PEL98) (MTS1 protein)
DE (Metastatic cell protein) (Fragment).
GN MTS1.
OS Mus spretus (Western white mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10096;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRET/BI;
RX MEDLINE=94319082; PubMed=8043949;
RA Ko M.S., Wang X., Horton J.H., Hagen M.D., Takahashi N., Maerzaki Y.,
RA Nadeau J.H.;
RT "Genetic mapping of 40 cDNA clones on the mouse genome by PCR.";
RT Mamm. Genome 5:349-355(1994).
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, AND MORE
CC SPECIFICALLY TO S-100/CABP LIKE PROTEINS.
DR EMBL; U05636; AAA61936.1; -.
DR PIR; I49407; I49407.
DR GO; GO:000509; F:calcium ion binding; IEA.
KW Calcium-binding; Placenta.
FT NON TER 1
SQ SEQUENCE 15 AA; 1824 MW; 31BBA50E3987741 CRC64;

Query Match 87.1%; Score 74; DB 11; Length 15;
Best Local Similarity 86.7%; Pred. No. 6.3e-06;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NEFFEGFDPKQPRK 15
   |||||
Db 1 NEFFEGCDKEPRK 15

RESULT 3
Q9M077 PRELIMINARY; PRT; 282 AA.
AC Q9M077;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN AT4G07710.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL161507; CAB7930.1; -.
DR PIR; A85076; A85076.
KW Hypothetical protein.
SQ SEQUENCE 282 AA; 31518 MW; 1A691C1C95991A03 CRC64;

Query Match 55.3%; Score 47; DB 10; Length 282;
Best Local Similarity 50.0%; Pred. No. 6.4;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 NEFFEGFDPKQPRK 14
   |||||
Db 17 NRYEGFDDAQTHR 30

RESULT 4
Q8I560 PRELIMINARY; PRT; 425 AA.
AC Q8I560;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN PFL1755W.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Bisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallom S.J., Suh B., Peterson J., Angiuoli S.,
RA Perteira M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum";
RL Nature 419:498-511(2002).
DR EMBL; AE014850; AAN36437.1; -.
DR InterPro; IPR008941; TPR-like.
KW Hypothetical protein.
SQ SEQUENCE 425 AA; 50761 MW; 1D9AACB2D7171791 CRC64;

Query Match 52.9%; Score 45; DB 5; Length 425;
Best Local Similarity 66.7%; Pred. No. 22;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 FEGFDPKQPRK 15
   |||||
Db 105 FENFPDKYKKK 116

RESULT 5
Q8M6G7 PRELIMINARY; PRT; 170 AA.
AC Q8M6G7;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN ORF60.
OS Sinorhizobium meliloti phage PBC5.
OC Viruses.
OX NCBI_TaxID=179237;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulmeister S.A., Krol J.E., Vorhoeelter F.-J., Skorupska A.M.,
RA Lotz W.;
RA "Sequence of the genome of Sinorhizobium meliloti bacteriophage
RT PBC5.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF448724; AAL49625.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR001387; HTH_3.
DR Pfam; PF01381; HTH_3; 1.
DR SMART; SM00530; HTH_XRE; 1.
KW Hypothetical protein.
SQ SEQUENCE 170 AA; 18934 MW; AF2F1A149B654FF1 CRC64;
```

```

Query Match          51.8%; Score 44; DB 9; Length 170;
Best Local Similarity 63.6%; Pred. No. 12;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 5 EGFPDKQPRKK 15
Db 112 EGLPDEQPRE 122

RESULT 6
ID Q8TPF9 PRELIMINARY; PRT; 262 AA.
AC Q8TPF9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DE 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DS Hypothetical protein MA1954.
GN MA1954.
OS Methanosarcina acetivorans.
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
OX NCBI_TaxID=2214;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C2A / ATCC 35395 / DSM 2834;
RX MEDLINE=21929760; PubMed=11932238;
RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA Allen N., Naylor J., Stange-Thomann N., Dearellano K., Johnson R.,
RA Linton L., McEwan P., McKernan K., Talanas J., Tirrell A., Ye W., A.M.,
RA Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA Metcalf W.W., Birren B.;
RT "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT and physiological diversity."
RL Genome Res. 12:532-542(2002).
DR EMBL; AE010878; AAM05357.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 262 AA; 29600 MW; D321454D57DD0FC8 CRC64;

Query Match          51.8%; Score 44; DB 17; Length 262;
Best Local Similarity 66.7%; Pred. No. 19;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 FEGFPDKQPRKK 15
Db 159 FAGPFMKYPRKK 170

RESULT 7
ID Q92YI1 PRELIMINARY; PRT; 370 AA.
AC Q92YI1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative ABC transporter, ATP-binding protein.
GN RA0899 OR SM1646.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,

```

```

RA Barloy-Hubler F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
DR EMBL; AE007276; AAK65557.1; -.
DR PIR; C95374; C95374.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. .; IEA.
DR GO; GO:0000166; F:nucleotide binding; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR GO; GO:0003593; AAA ATPase.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00005; ABC tran; 1.
DR ProDom; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA; 1.
DR PROSITE; PS00211; ABC TRANSPORTER 1; 1.
DR PROSITE; PS00893; ABC TRANSPORTER 2; 1.
KW ATP-binding; Plasmid; Hypothetical protein; Complete proteome.
SQ SEQUENCE 370 AA; 40728 MW; ABD9738616E1007A CRC64;

Query Match          51.8%; Score 44; DB 16; Length 370;
Best Local Similarity 63.6%; Pred. No. 28;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 FEGFPDKQPRKK 14
Db 17 FRGLPNRQPRK 27

RESULT 8
ID Q8EWA0 PRELIMINARY; PRT; 282 AA.
AC Q8EWA0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Conserved hypothetical protein.
GN MYPE3040.
OS Mycoplasma penetrans.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=28227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HF-2;
RX MEDLINE=22354719; PubMed=12466555;
RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT "The complete genomic sequence of Mycoplasma penetrans, an
RT intracellular bacterial pathogen in humans."
RL Nucleic Acids Res. 30:5293-5300(2002).
DR EMBL; AP004171; BAC44096.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 282 AA; 33070 MW; 113CE5FB9A0E139A CRC64;

Query Match          50.6%; Score 43; DB 16; Length 282;
Best Local Similarity 53.3%; Pred. No. 31;
Matches 8; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 NEFFEGFPDKQPRKK 15
Db 110 NEIFGGIPDGEFVNK 124

RESULT 9
ID Q9CQ51 PRELIMINARY; PRT; 330 AA.
AC Q9CQ51;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)

```

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE DNA segment, Chr 16, Wayne state University 83, expressed.
 GN TWSTNB OR D16WSU83E.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=Testis, and Embryonic stem cells;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuoka H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.P.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL; AK016638; BAB30352.1; -;
 DR EMBL; AK010829; BAB27210.1; -;
 DR MGD; MGI:106292; Twistnb.
 DR InterPro; IPR008994; Nucleic acid OB.
 DR Pfam; PF05156; RNA_pol_Rp43; 1.
 SQ SEQUENCE 330 AA; 36721 MW; C68996895950DA73 CRC64;

 Query Match 50.6%; Score 43; DB 11; Length 330;
 Best Local Similarity 53.3%; Pred. No. 37;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

 QY 1 NEFFEGFPDQPKK 15
 Db 271 NDFEFPKPKKKKK 285

 RESULT 10
 ID Q8Q0I4 PRELIMINARY; PRT; 378 AA.
 AC Q8Q0I4;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DE DNA integration/recombination/inversion protein.
 GN M00152.
 OS Methanosarcina mazei (Methanosarcina frisia).
 OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;
 OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.
 OX NCBI_TaxID=2209;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Goel / Go1 / ATCC BAA-199 / DSM 3647 / OCM 88;
 RX MEDLINE=22120827; PubMed=12125824;
 RA Deppenmeier U., Johann A., Hartsch T., Merkl R., Schmitz R.A.,
 RA Martinez-Arias R., Henne A., Wietzer A., Baumer S., Jacobi C.,
 RA Brueggemann H., Lienard T., Christmann A., Roemcke M., Steckel S.,
 RA Bhattacharyya A., Lykidis A., Overbeek R., Klenk H.-P., Gunsalus R.P.,
 RA Fritz H.-J., Gottschalk G.;
 RT "The genome of Methanosarcina mazei: evidence for lateral gene
 transfer between Bacteria and Archaea.";

RL J. Mol. Microbiol. Biotechnol. 4:453-461(2002).
 DR EMBL; A5013236; AAM29848.1; -;
 DR GO; GO:0003677; F:DNA binding; IEA.
 DR GO; GO:0015074; P:DNA integration; IEA.
 DR GO; GO:0006310; P:DNA recombination; IEA.
 DR InterPro; IPR002104; Phage integrase.
 DR InterPro; IPR004107; Phage_integr_N.
 DR Pfam; PF00589; Phage_integrase; 1.
 DR Pfam; PF02899; Phage_integr_N; 1.
 KW Complete proteome.
 SQ SEQUENCE 378 AA; 43233 MW; C2D65995BC4B2B59 CRC64;

 Query Match 50.6%; Score 43; DB 17; Length 378;
 Best Local Similarity 53.8%; Pred. No. 43;
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

 QY 1 NEFFEGFPDQPKR 13
 Db 93 NDFFTGKSKQPK 105

 RESULT 11
 ID Q9FN73 PRELIMINARY; PRT; 758 AA.
 AC Q9FN73;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Gb|AA09076.1.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Columbia;
 RX MEDLINE=98069011; PubMed=9405937;
 RA Kotani H., Nakamura Y., Sato S., Kaneko T., Asamizu E., Miyajima N.,
 RA Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. II.
 RT Sequence features of the regions of 1,044,062 bp covered by thirteen
 RT physically assigned pl clones.";
 RL DNA Res. 4:291-300(1997).
 DR EMBL; AB006706; BAB09578.1; -;
 DR InterPro; IPR001876; Znf_RanGDP.
 DR Pfam; PF00641; zf-RanBP; 2.
 DR SMART; SM00547; Znf_RBZ; 2.
 DR PROSITE; PS01358; ZF_RANBP2_1; 2.
 DR PROSITE; PS0159; ZF_RANBP2_2; 2.
 SQ SEQUENCE 758 AA; 85944 MW; 5659AABF750BF8DE CRC64;

 Query Match 50.6%; Score 43; DB 10; Length 758;
 Best Local Similarity 70.0%; Pred. No. 89;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

 QY 1 NEFFEGFPDK 10
 Db 119 SEFFFGFPDE 128

 RESULT 12
 ID Q8S9K3 PRELIMINARY; PRT; 758 AA.
 AC Q8S9K3;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE AT5g17790/VVA3.140.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.

```

OX NCBI_TaxID=3702;
RN [1]
RN SEQUENCE FROM N.A.
RA Shinn P., Chen H., Cheuk R., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
RA Kawai J., Lam J., Lee J.M., Lin J., Miranda M., Narusaka M., Narusaka M.,
RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
RA Ecker J.R.;
RT "Arabidopsis cDNA clones.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RA Cheuk R., Chen H., Kim C.J., Shinn P., Bowser L., Carninci P.,
RA Chan M.M., Chang C.H., Dale J.M., Hayashizaki Y., Hsuan V.W.,
RA Ishida J., Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Lam B.,
RA Lee J.M., Lin J., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Quach H.L., Sakurai T., Satou M., Seki M., Southwick A., Tang C.C.,
RA Toriumi M., Wallender E.K., Wong C., Wu H.C., Yamada K., Yu G.,
RA Yuan S., Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;
RT "Arabidopsis ORF clones.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY075645; AAL77653.1; -.
DR EMBL; BT002233; AAN72244.1; -.
DR InterPro; IPR001876; Znf_RanGDP.
DR Pfam; PF00641; zf-RanBP; 2.
DR SMART; SM00547; Znf_RBZ; 2.
DR PROSITE; PS01358; ZF_RANBP2_1; 2.
DR PROSITE; PS01199; ZF_RANBP2_2; 2.
SQ SEQUENCE 758 AA; 86013 MW; 4CE8000F741AF9CE CRC64;

Query Match
Best Local Similarity 50.6%; Score 43; DB 10; Length 758;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NEFFEGFPDK 10
Db 119 SEFFGFPDE 128
:|||||:
:|||||:

RESULT 13
Q80520
ID Q80520 PRELIMINARY; PRT; 997 AA.
AC Q80520
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Iibonucleotide reductase large subunit.
GN UL39.
OS Simian herpes B virus (Cercopithecid herpesvirus 1) (Shbv).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10325;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=E2490;
RX MEDLINE=22607624; PubMed=12721804;
RA Ohsawa K., Black D.H., Sato H., Rogers K., Eberle R.;
RA "Sequence and genetic arrangement of the UL region of the monkey B
RT virus (Cercopithecine herpesvirus 1) genome and comparison with the UL
RL region of other primate herpesviruses.";
RL Arch. Virol. 148:989-997 (2003).
DR EMBL; AB096160; BAC58079.1; -.
DR GO; GO:0005971; C:ribonucleoside-diphosphate reductase complex; IEA.
DR GO; GO:0004748; F:ribonucleoside-diphosphate reductase activity; IEA.
DR GO; GO:0006260; P:DNA replication; IEA.
DR InterPro; IPR000788; Ribonucleo.
DR Pfam; PF00317; ribonuc_red_lg; 1.
DR Pfam; PF02867; ribonuc_red_lgc; 1.
DR PRINTS; PR01183; RIBORDASEM1.
DR PROSITE; PS00089; RIBORED_LARGE; 1.
SQ SEQUENCE 997 AA; 111439 MW; 7544D8DF0C430F21 CRC64;

```

```

Query Match
Best Local Similarity 50.6%; Score 43; DB 12; Length 997;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 FEGFPDKQPR 13
Db 792 WEGFPDARPR 801
:|||||:
:|||||:

RESULT 14
Q7TSD7
ID Q7TSD7 PRELIMINARY; PRT; 997 AA.
AC Q7TSD7
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Large subunit of ribonucleotide reductase.
GN U139.
OS Simian herpes B virus (Cercopithecid herpesvirus 1) (Shbv).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Simplexvirus.
OX NCBI_TaxID=10325;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=E2490;
RX MEDLINE=22628476; PubMed=12743273;
RA Perelygina L., Zhu L., Zurkuhlen H., Mills R., Borodovsky M.,
RA Hilliard J.K.;
RT "Complete Sequence and Comparative Analysis of the Genome of Herpes B
RT Virus (Cercopithecine Herpesvirus 1) from a Rhesus Monkey.";
RL J. Virol. 77:6167-6177 (2003).
DR EMBL; AF533768; AAP41457.1; -.
SQ SEQUENCE 997 AA; 111425 MW; 99A0B4220D1A6F1B CRC64;

Query Match
Best Local Similarity 50.6%; Score 43; DB 12; Length 997;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 4 FEGFPDKQPR 13
Db 792 WEGFPDARPR 801
:|||||:
:|||||:

RESULT 15
Q98BN5
ID Q98BN5 PRELIMINARY; PRT; 625 AA.
AC Q98BN5
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Peptide ABC transporter, periplasmic peptide-binding
DE protein.
GN MLI5494.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RL Mesorhizobium loti.";
RL DNA Res. 7:331-338 (2000).
DR EMBL; AP003006; BAB51937.1; -.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.

```

DR InterPro; IPR000914; SBP_bac_5.
DR Pfam; PF00496; SBP_bac_5; 1.
KW Complete proteome.
SQ SEQUENCE 625 AA; 70446 MW; 7672D95717853DB8 CRC64;
Query Match 49.4%; Score 42; DB 16; Length 625;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 6 GFDDKQP 12
|||
Db 591 GFDDKQP 597
Search completed: September 9, 2004, 10:58:22
Job time : 26.9677 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: September 9, 2004, 10:32:57 ; Search time 3.69032 Seconds
(without alignments)
183.429 Million cell updates/sec

Title: US-10-067-618-6

Perfect score: 65

Sequence: 1 LPSFLGKRTDEAA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	100.0	101	1	P26447 homo sapien
2	65	100.0	101	1	P07091 mus musculus
3	62	95.4	101	1	P05942 rattus norv
4	61	93.8	100	1	P25466 bos taurus
5	44	67.7	221	1	Q42376 zea mays (m
6	42	64.6	97	1	P10462 bos taurus
7	42	64.6	97	1	P29034 homo sapien
8	40	61.5	637	1	Q985Y3 rhizobium l
9	38	58.5	644	1	O52027 halobacteri
10	38	58.5	2109	1	P03523 vesicular s
11	37	56.9	87	1	P00114 synecococc
12	37	56.9	112	1	P56534 synecococc
13	37	56.9	112	1	Q9f119 synecococc
14	37	56.9	280	1	Q9ju48 n 2-dehydro
15	37	56.9	284	1	P57853 p 2-dehydro
16	37	56.9	384	1	P57308 buchnera ap
17	36	55.4	84	1	Q10329 orgyia pseu
18	36	55.4	280	1	Q881x0 p 2-dehydro
19	36	55.4	280	1	Q9jz55 n 2-dehydro
20	36	55.4	281	1	Q88mg0 p 2-dehydro
21	36	55.4	281	1	Q9zfik p 2-dehydro
22	36	55.4	281	1	Q886m4 pseudomonas
23	36	55.4	282	1	Q9ear9 s 2-dehydro
24	36	55.4	283	1	Q9kq29 vibrio chol
25	36	55.4	283	1	Q87rn0 v 2-dehydro
26	36	55.4	284	1	Q8xds7 e 2-dehydro
27	36	55.4	284	1	Q9fzh8 escherichia
28	36	55.4	284	1	P17579 escherichia
29	36	55.4	284	1	P45251 h 2-dehydro
30	36	55.4	284	1	P95514 p 2-dehydro
31	36	55.4	284	1	Q8xgr9 salmonella
32	36	55.4	284	1	Q8dfg3 vibrio vuln
33	36	55.4	284	1	Q8ze4 yersinia pe

34	36	55.4	286	1	KD5A_ACTEL
35	36	55.4	384	1	HYMA_EMENI
36	36	55.4	389	1	METK_NEIMA
37	36	55.4	389	1	METK_NEIMB
38	36	55.4	457	1	BAG4_HUMAN
39	36	55.4	457	1	BAG4_MOUSE
40	36	55.4	880	1	SYA_ANASP
41	36	55.4	2109	1	RRPL_VSVJH
42	36	55.4	2109	1	RRPL_VSVJO
43	35	53.8	228	1	ZFPI_ARATH
44	35	53.8	316	1	P567_ARATH
45	35	53.8	384	1	FISZ_EUCAP

068662 a 2-dehydro
O60032 emericella
Q9jvv6 neisseria m
Q9jv09 neisseria m
Q95429 homo sapien
Q8c161 mus musculus
Q8yud4 anabaena sp
P13615 vesicular s
P16379 vesicular s
Q42485 arabidopsis
Q91vl2 arabidopsis
O51929 buchnera ap

RESULT 1
S104_HUMAN
ID S104_HUMAN STANDARD; PRT; 101 AA.
AC P26447;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Placental calcium-binding protein (Calvasculin) (S100 calcium-binding protein A4) (MTS1 protein).
DE GN S100A4 OR CAPL OR MTS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=93041710; PubMed=1384693;
RA Engelkamp D., Schaefer B.W., Erne P., Heizmann C.W.;
RT "S100 alpha, CAPL, and CACY: molecular cloning and expression analysis of three calcium-binding proteins from human heart.";
RL Biochemistry 31:10258-10264(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93342029; PubMed=8341667;
RA Engelkamp D., Schaefer B., Mattel M.-G., Erne P., Heizmann C.W.;
RT "Six S100 genes are clustered on human chromosome 1q21: identification of two genes coding for the two previously unreported calcium-binding proteins S100D and S100E.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:6547-6551(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=93028421; PubMed=1329089;
RA Tulchinsky E.M., Ford H.L., Kramerov D., Reshetnyak E., Grigorian M., Zain S., Lukanidin E.;
RT "Transcriptional analysis of the mts1 gene with specific reference to 5' flanking sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:9146-9150(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix, and Prostate;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Raheij J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

ALIGNMENTS

```

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M80563; AAA51920.1; -
DR EMBL; Z18950; CAA79474.1; -
DR EMBL; Z33457; CAA83880.1; -
DR EMBL; BC000838; AAH00838.1; -
DR EMBL; BC016300; AAH16300.1; -
DR PIR; A48219; A48219.
DR PDB; 1M31; 30-OCT-02.
DR SWISS-2DPAGE; P26447; HUMAN.
DR Genew; HGNC:10494; S100A4.
DR MIM; 114210; -.
DR GO; GO:0005509; F.calcium ion binding; TAS.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CaBP; 1.
DR Calcium-binding; 3D-structure.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SQ SEQUENCE 101 AA; 11728 MW; 28602B7B07EDB562 CRC64;

Query Match 100.0%; Score 65; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
DB 42 LPSFLGKRTDEAA 54

RESULT 2
ID S104_MOUSE STANDARD; PRT; 101 AA.
AC F07091; P20066;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Placental calcium-binding protein (18A2) (PEL98) (MTS1 protein)
DE (Metastatic cell protein).
GN S100A4 OR CAPL OR MTS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=87316927; PubMed=3628004;
RA Jackson-Grusby L.L., Swiergiel J., Linzer D.I.H.;
RT "A growth-related mRNA in cultured mouse cells encodes a placental
RT calcium binding protein."

```

```

Nucleic Acids Res. 15:6677-6690 (1987).
[2]
SEQUENCE FROM N.A.
STRAIN=BALB/c;
MEDLINE=88198109; PubMed=3162911;
RA Goto K., Endo H., Fujiyoshi T.;
RT "Cloning of the sequences expressed abundantly in established cell
RT lines: identification of a cDNA clone highly homologous to S-100, a
RT calcium binding protein.";
RL J. Biochem. 103:48-53 (1988).
[3]
SEQUENCE FROM N.A.
MEDLINE=89378739; PubMed=2550322;
RA Ebralidze A., Tulchinsky E., Grigorian M., Afanasyeva A., Senin V.,
RA Revazova E., Lukanidin E.;
RT "Isolation and characterization of a gene specifically expressed in
RT different metastatic cells and whose deduced gene product has a high
RT degree of homology to a Ca2+-binding protein family.";
RL Genes Dev. 3:1086-1093 (1989).
[4]
SEQUENCE FROM N.A.
TISSUE=Liver;
RX MEDLINE=90236313; PubMed=2332170;
RA Tulchinsky E.M., Grigorian M.S., Ebralidze A.K., Milshina N.I.,
RA Lukanidin E.M.;
RT "Structure of gene mts1, transcribed in metastatic mouse tumor
RT cells";
RL Gene 87:219-223 (1990).
[5]
SEQUENCE OF 1-54 FROM N.A.
MEDLINE=93141279; PubMed=8423998;
RA Tulchinsky E., Kramerov D., Ford H.L., Reshetnyak E., Lukanidin E.,
RA Zain S.;
RT "Characterization of a positive regulatory element in the mts1 gene.";
RL Oncogene 8:79-86 (1993).
CC -!- TISSUE SPECIFICITY: Specifically expressed in different metastatic
CC cells.
CC -!- INDUCTION: The mRNA coding for this protein increases in
CC abundance after serum stimulation of quiescent mouse fibroblasts.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X05835; CAA29282.1; -
DR EMBL; D00208; BAA00148.1; -
DR EMBL; M36578; AAA39749.1; -
DR EMBL; M36579; AAA39750.1; -
DR EMBL; X16190; CAA34316.1; -
DR EMBL; X16094; CAA34224.1; -
DR PIR; S06207; S06207.
DR HSP; P30801; 1A03.
DR MGD; MGI:1330282; S100a4.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CaBP; 1.
DR Calcium-binding; Placenta.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT CONFLICT 47 47 GVSXKFNQ (IN REF. 5).
SQ SEQUENCE 101 AA; 11721 MW; 2302254B67A4C873 CRC64;

Query Match 100.0%; Score 65; DB 1; Length 101;

```


Best Local Similarity 100.0%; Pred. NO. 2e-05; Mismatches 0; Indels 0; Gaps 0;
Matches 13; Conservative 0;

QY 1 LPSFLGKRTDEAA 13
|||||:|||||
Db 42 LPSFLGKRTDEAA 54

RESULT 3

ID S104 RAT STANDARD; PRT; 101 AA.
AC P05942;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Placental calcium-binding protein (Nerve growth factor induced protein 42A) (P9K).
GN S100A4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88125019; PubMed=3422491;
RA Maslakowski P., Shooter E.M.;
RT "Nerve growth factor induces the genes for two proteins related to a family of calcium-binding proteins in PC12 cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:1277-1281(1988).
RN [2]

RP SEQUENCE FROM N.A.
RX MEDLINE=88118907; PubMed=3430604;
RA Barraclough R., Savin J., Dube S.K., Rudland P.S.;
RT "Molecular cloning and sequence of the gene for p9Ka. A cultured myoepithelial cell protein with strong homology to S-100, a calcium-binding protein.";
RL J. Mol. Biol. 198:13-20(1987).
RN [3]

RP SEQUENCE OF 3-101 FROM N.A.
RX MEDLINE=92158347; PubMed=1741158;
RA de Vonge M.W., Mukherjee B.B.;

RT "Transformation of normal rat kidney cells by v-K-ras enhances expression of transin 2 and an S-100-related calcium-binding protein.";

RL Oncogene 7:109-119(1992).

CC -1- INDUCTION: By nerve growth factor.

CC -1- SIMILARITY: Belongs to the S-100 family.

CC -1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC ENBL; X06916; CAA30014.1; --

DR ENBL; J03628; AAA42098.1; --

DR ENBL; X64024; --; NOT_ANNOTATED_CDS.

DR ENBL; X64023; --; NOT_ANNOTATED_CDS.

DR PIR; S01759; S01759.

DR HSSP; P30801; 1A03.

DR InterPro; IPR001751; CaBP_S100.

DR InterPro; IPR002045; EF-hand.

DR Pfam; PF00036; ehand; 1.

DR Pfam; PF01023; S_100; 1.

DR ProDom; PD003407; CaBP_S100; 1.

DR PROSITE; PS00018; EF_HAND; 1.

DR PROSITE; PS00303; S100_CaBP; 1.

KW Calcium-binding.

FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).

FT CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).

SQ SEQUENCE 101 AA; 11776 MW; EA0619CEB4F487C1 CRC64;

Query Match 95.4%; Score 62; DB 1; Length 101;
Best Local Similarity 92.3%; Pred. No. 7.6e-05;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
|||||:|||||
Db 42 LPSFLGRTDEAA 54

RESULT 4

ID S104 BOVIN STANDARD; PRT; 100 AA.
AC P35466;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Placental calcium-binding protein homolog.

GN S100A4 OR CAPL.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP SEQUENCE.

RC TISSUE=Retina;

RX MEDLINE=94164991; PubMed=8119967;

RA Polans A.S., Palczewski K., Asson-Batres M.A., Ohguro H., Witowska D.,

RA Haley T.L., Baizer L., Crabb J.W.;

RT "Purification and primary structure of Cap1, an S-100-related calcium-binding protein isolated from bovine retina.";

RL J. Biol. Chem. 269:6233-6240(1994).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=98430466; PubMed=9759666;

RA Duarte W.R., Kasugai S., Iimura T., Oida S., Takenaga K., Ohya K.,

RA Ishikawa I.;

RT "CDNA cloning of S100 calcium-binding proteins from bovine periodontal ligament and their expression in oral tissues.";

RL J. Dent. Res. 77:1694-1699(1998).

CC -1- SIMILARITY: Belongs to the S-100 family.

CC -1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC ENBL; D89056; BAA13754.1; --

DR PIR; A53217; A53217.

DR HSSP; P30801; 1A03.

DR InterPro; IPR001751; CaBP_S100.

DR InterPro; IPR002048; EF-hand.

DR Pfam; PF00036; ehand; 1.

DR Pfam; PF01023; S_100; 1.

DR ProDom; PD003407; CaBP_S100; 1.

DR PROSITE; PS00018; EF_HAND; 1.

DR PROSITE; PS00303; S100_CaBP; 1.

KW Calcium-binding; Acetylation.

FT INIT_MET 0 0 ACETYLTATION.

FT MOD_RES 1 1

FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).

FT CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).

SQ SEQUENCE 100 AA; 11675 MW; DFFCAA7561D5EFA4 CRC64;

Query Match 93.8%; Score 61; DB 1; Length 100;
Best Local Similarity 92.3%; Pred. No. 0.00012;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
DB 41 LPSFLGKRTDETA 53

RESULT 5

LEA3 MAIZE
ID LEA3 MAIZE STANDARD; PRT; 221 AA.

AC Q42376;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Late embryogenesis abundant protein, group 3 (LEA).
GN MGL3.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC PACCAD clade; Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=95357439; PubMed=7630968;
RA White C.N., Rivin C.J.;
RT "Sequence and regulation of a late embryogenesis abundant group 3
RT protein of maize.";
RL Plant Physiol. 108:1337-1338(1995).
CC -!- SIMILARITY: Belongs to the LEA type 1 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U05226; AAA83402.1; -;
DR EMBL; D26552; BAA05550.1; -;
DR EMBL; Z29512; CAA82632.1; -;
DR PIR; S41387; S41387; LEA.
DR InterPro; IPR004238; LEA.
DR Pfam; PF02987; LEA; 2.
KW Repeat; Seed; Seed embryo.
SQ SEQUENCE 221 AA; 22759 MW; C649EA0B1D0DB169 CRC64;

Query Match 67.7%; Score 44; DB 1; Length 221;
Best Local Similarity 80.0%; Pred. No. 0.5;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 SPFLGKRTDEA 12
DB 68 SYLGKRTDEA 77

RESULT 6

S102 BOVIN
ID S102 BOVIN STANDARD; PRT; 97 AA.

AC P10452;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE S100 calcium-binding protein A2 (S-100L protein).
GN S100A2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP Wicki R., Franz C., Scholl F.A., Heizmann C.W., Schaefer B.W.;
RT "repression of the candidate tumor suppressor gene S100A2 in breast
RT cancer is mediated by site-specific hypermethylation.";
RL Cell Calcium 22:243-254(1997).

DR Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
DR Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
DR Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
DR Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heif F.,
DR Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
DR Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
DR Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
DR Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
DR Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
DR Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

DR Glemey J.R. Jr., Kindy M.S., Zokas L.;
RT "Isolation of a new member of the S100 protein family: amino acid
RT sequence, tissue, and subcellular distribution.";
RL J. Cell Biol. 108:569-578(1989).
CC -!- SUBUNIT: Homomultimeric (Probable).
CC -!- MISCELLANEOUS: This protein binds two calcium ions.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
DR PIR; A30129; A30129.
DR HSP; P30801; 1A03.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CaBP; 1.
KW Calcium-binding.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SQ SEQUENCE 97 AA; 10893 MW; 55ACC9F60CF9C9E6 CRC64;

Query Match 64.6%; Score 42; DB 1; Length 97;
Best Local Similarity 63.6%; Pred. No. 0.53;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDE 11
DB 42 LPSFVGKRVDE 52

RESULT 7

S102 HUMAN
ID S102 HUMAN STANDARD; PRT; 97 AA.

AC P29034; O00266; Q9BU83;
DT 01-DEC-1992 (Rel. 24, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE S100 calcium-binding protein A2 (S-100L protein) (CAN19).
GN S100A2 OR S100L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=92196147; PubMed=1372446;
RA Lee S.W., Tomasetto C., Swisshelm K., Keyomarsi K., Sager R.;
RT "Down-regulation of a member of the S100 gene family in mammary
RT carcinoma cells and reexpression by azadeoxycytidine treatment.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2504-2508(1992).

DR Wicki R., Franz C., Scholl F.A., Heizmann C.W., Schaefer B.W.;
RT "repression of the candidate tumor suppressor gene S100A2 in breast
RT cancer is mediated by site-specific hypermethylation.";
RL Cell Calcium 22:243-254(1997).

DR Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
DR Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
DR Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
DR Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heif F.,
DR Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
DR Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
DR Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
DR Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
DR Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
DR Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences.";
PL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [4]
RP SEQUENCE OF 36-39 AND 41-48.
RC TISSUE=Keratinocytes;
RX MEDLINE=93162043; PubMed=1286667;
RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
RA Vandekerckhove J.;
RT "Microsequences of 145 proteins recorded in the two-dimensional gel
protein database of normal human epidermal keratinocytes.";
RL Electrophoresis 13:960-969 (1992).
CC -!- FUNCTION: May act as a modulator against excess calcium
CC accumulation in normal human mammary epithelial cells. May also
CC play a role in suppressing tumor cell growth.
CC -!- TISSUE SPECIFICITY: A subset of epithelial cells amongst which
CC normal human mammary epithelial cells and keratinocytes.
CC -!- DEVELOPMENTAL STAGE: Preferentially expressed in normal human
CC mammary epithelial cells as opposed to tumor-derived ones. The
CC level of S100L was shown to correlate inversely with tumor
CC progression.
CC -!- INDUCTION: By growth factors in early G1 phase and probably by
CC cell-cycle regulation in S phase. DNA methylation probably plays
CC a direct negative role in suppressing S100L gene expression in
CC tumor cells.
CC -!- MISCELLANEOUS: This protein binds two calcium ions (By
CC similarity).
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M87068; -; NOT ANNOTATED_CDS.
CC EMBL; Y07755; CAA69033.1; -;
CC EMBL; BC002829; AAH02829.2; -;
CC HSP; P30801; IAA03.
CC Aarhus/Ghent-2DPAGE; 9027; IEF.
CC Aarhus/Ghent-2DPAGE; 9051; IEF.
CC Genew; HGNC:10492; S100A2.
CC MTM; 176993; -;
CC GO; GO:0005509; F:calcium ion binding; NAS.
CC InterPro; IPR001751; CaBP_S100.
CC InterPro; IPR002048; EF-hand.
CC Pfam; PF01023; S100; 1.
CC ProDom; PD003407; CaBP_S100; 1.
CC ProDom; PD000012; EF-hand; 1.
CC PROSITE; PS00018; EF_HAND; 1.
CC PROSITE; PS00303; S100_CaBP; 1.
CC Calcium-binding.
CC CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
CC CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
CC FT CONFLICT 61 61 S -> N (IN REF. 1).
CC FT CONFLICT 61 61 S -> N (IN REF. 1).
CC SQ SEQUENCE 97 AA; 10966 MW; CFB06CD8DC08D250 CRC64;

Query Match 64.6%; Score 42; DB 1; Length 97;
Best Local Similarity 63.6%; Pred. No. 0.53;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 LPSFLGKRTDE 11
|||||:::|

Db 42 LPSFVGKVDK 52

RESULT 8
DXS_RHILO
ID DXS_RHILO STANDARD; PRT; 637 AA.
AC Q985Y3;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE 1-deoxy-D-xylulose 5-phosphate synthase (EC 2.2.1.7) (1-
DE deoxyxylulose-5-phosphate synthase) (DXP synthase) (DXPS).
GN DXS OR MLR7474.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338 (2000).
CC -!- FUNCTION: Catalyzes the acylain condensation reaction between C
CC atoms 2 and 3 of pyruvate and glyceraldehyde 3-phosphate to yield
CC 1-deoxy-D-xylulose-5-phosphate (DXP) (By similarity).
CC -!- CATALYTIC ACTIVITY: Pyruvate + D-glyceraldehyde 3-phosphate = 1-
CC deoxy-D-xylulose 5-phosphate + CO(2).
CC -!- COFACTOR: Binds 1 thiamine pyrophosphate per subunit (By
CC similarity).
CC -!- PATHWAY: Nonaevonate terpenoid biosynthesis pathway; first step.
CC -!- PATHWAY: Biosynthetic pathway to thiamine and pyridoxol; first
CC step.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SIMILARITY: Belongs to the transketolase family. DXPS subfamily.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP003011; BAB53570.1; -;
CC HAMAP; MF 00315; -; 1.
CC InterPro; IPR005477; Dxs synth.
CC InterPro; IPR009014; Transketolase_C.
CC InterPro; IPR005476; Transketolase_C.
CC InterPro; IPR005475; Transketolase_CR.
CC InterPro; IPR005474; Transketolase_N.
CC Pfam; PF02779; transket pyr; 1.
CC Pfam; PF02780; transketolase_C; 1.
CC TIGRfams; TIGR00204; dxs; 1.
CC PROSITE; PS00801; TRANSKETOLASE_1; 1.
CC PROSITE; PS00802; TRANSKETOLASE_2; 1.
CC Transfrase; Flavoprotein; Thiamine pyrophosphate;
CC Isoprene biosynthesis; Thiamine biosynthesis; Complete proteome.
CC SQ SEQUENCE 637 AA; 67946 MW; 3A396E1DA6E50570 CRC64;

Query Match 61.5%; Score 40; DB 1; Length 637;
Best Local Similarity 66.7%; Pred. No. 8.7;
Matches 8; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 LPSFLGKRTDE 12
|||||:::|

Db 214 LPSYLGKRDRA 225

```
RESULT 9
ARSA_HALN1 STANDARD; PRT; 644 AA.
AC 052027;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative arsenical pump-driving ATPase (EC 3.6.3.16) (Arsenite-
DE translocating ATPase) (Arsenical resistance ATPase) (Arsenite-
DE transporting ATPase).
GN ARSA OR ARSA2 OR VNG5180G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OG Plasmid pNRC100.
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99063795; PubMed=9847077;
RA Ng W.V., Clufo S.A., Smith T.M., Bumgarner R.E., Baskin D., Faust J.,
RA Hall B., Loretz C., Seto J., Slagel J., Hood L., Dassarma S.;
RT "Snapshot of a large dynamic replicon in a halophilic archaeon:
RT megaplasmid or minichromosome?";
RL Genome Res. 8:1131-1141(1998).
CC -!- FUNCTION: Anion-transporting ATPase. Catalyzes the extrusion of
CC arsenite (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + H(2)O + arsenite(in) = ADP + phosphate +
CC arsenite(out).
CC -!- SIMILARITY: Belongs to the arsa ATPase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X00939; CAA25453.1; --
CC PIR; A04113; ZLVN.
CC InterPro; IPR007098; RNA_pol_monon.
CC InterPro; IPR002877; RrmJ_FtsJ.
CC InterPro; IPR001016; Viral_RNA_pol_L.
CC Pfam; PF01728; FtsJ; 1.
CC Pfam; PF00946; Paramyx_RNA_pol; 1.
CC Transferase; RNA-directed RNA polymerase.
KW SEQUENCE 2109 AA; 240880 MW; 011EC2C0B5967C27 CRC64;
SQ
Query Match 58.5%; Score 38; DB 1; Length 2109;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
QY 1 LPSFLGKRTDEA 12
DB 1149 LPAYLGSKTSES 1160
RESULT 11
CYC6_SYNL1 STANDARD; PRT; 87 AA.
AC P00114;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cytochrome C6 (Soluble cytochrome f) (Cytochrome c553) (Cytochrome c-
DE 553).
DE GN PETJ.
OS Synechococcus lividus.
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=33070;
RN [1]
RP SEQUENCE.
RA Borden D., Margoliash E.;
RL Submitted (DEC-1979) to the PIR data bank.
CC -!- FUNCTION: Functions as an electron carrier between membrane-bound
CC cytochrome b6f and photosystem I in oxygenic photosynthesis (By
CC similarity).
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
CC -!- PTM: Binds 1 heme group per molecule.
CC -!- SIMILARITY: Belongs to the cytochrome c family. PetJ subfamily.
CC PIR; A00106; CCYC6L.
CC HSP; P56534; IC6S.
CC HAMAP; MF_00594; -.
CC InterPro; IPR003088; Cyt_CI.
CC
```

```
RESULT 10
RRPL_VSVSJ STANDARD; PRT; 2109 AA.
AC P03523;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 26-FEB-2003 (Rel. 41, Last annotation update)
DE RNA polymerase beta subunit (EC 2.7.7.48) (Large structural protein)
DE (L protein).
GN L.
OS Vesicular stomatitis virus (strain San Juan).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Rhabdoviridae; Vesiculovirus.
CC
```

DR InterPro; IPR008168; Cyt C1C.
 DR InterPro; IPR000345; CytC_heme_BS.
 DR Pfam; PF00034: cytochrome c; 1.
 DR PRINTS; PR00605; CYTCHROME.C.
 DR PROSITE; PS00190; CYTOCHROME C; 1.
 KW Electron transport; Photosynthesis; Heme.
 FT BINDING 14 14 HEME (COVALENT).
 FT BINDING 17 17 HEME (COVALENT).
 FT METAL 18 18 IRON (HEME AXIAL LIGAND).
 FT METAL 58 58 IRON (HEME AXIAL LIGAND).
 SQ SEQUENCE 87 AA; 9129 MW; 37713EC6405EBEE CRC64;

Query Match 56.9%; Score 37; DB 1; Length 87;
 Best Local Similarity 54.5%; Pred. No. 4.4;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDE 11
 Db 58 MPAFAGRLTDE 68

RESULT 12
 ID CYC6 SYNEL STANDARD; PRT; 112 AA.
 AC P56534; Q9F1M0;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cytochrome C6 precursor (Soluble cytochrome f) (Cytochrome c553)
 DE (Cytochrome c-553).
 GN PETU OR TLL1283.
 OS Synecococcus elongatus (Thermosynechococcus elongatus), and
 OS Synecococcus elongatus naegeli.
 OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
 OX NCBI_TaxID=32046, 1141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.elongatus; STRAIN=BP-1;
 RA Katoh H., Itoh S., Shen J.-R., Ikeuchi M.;
 RT "Cloning of the genes for cytochrome c550 and a c550-like protein from
 RL the thermophilic cyanobacterium Synecococcus elongatus.";
 RN Plant Cell Physiol. 40:89-99(1999).
 [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.elongatus; STRAIN=BP-1;
 RX MEDLINE=2225144; PubMed=12240834;
 RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
 RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
 RA Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the thermophilic cyanobacterium
 RL Thermosynechococcus elongatus BP-1.";
 RN DNA Res. 9:123-130(2002).
 [3]
 RP STRUCTURE BY NMR OF 26-112.
 RC SPECIES=S.e.naegeli;
 RX MEDLINE=9809047; PubMed=9427738;
 RA Beissinger M., Sticht H., Sutter M., Eichart A., Haehnel W.,
 RA Roesch P.;
 RT "Solution structure of cytochrome c6 from the thermophilic
 RL cyanobacterium Synecococcus elongatus.";
 RN EMBO J. 17:27-36(1998).
 CC -1- FUNCTION: Functions as an electron carrier between membrane-bound
 CC cytochrome b6f and photosystem I in oxygenic photosynthesis.
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- PTM: Binds 1 heme group per molecule.
 CC -1- SIMILARITY: Belongs to the cytochrome c family. PetJ subfamily.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

DR EMBL; AB052597; BAB20061.1; --
 DR EMBL; AP005373; BAC08835.1; --
 DR FDB; 1C6S; 08-APR-98.
 DR HAMAP; MF_00594; --; 1.
 DR InterPro; IPR003088; Cyt C1.
 DR InterPro; IPR008168; Cyt_C1C.
 DR InterPro; IPR000345; CytC_heme_BS.
 DR Pfam; PF00034: cytochrome c; 1.
 DR PRINTS; PR00605; CYTCHROME.C.
 DR PROSITE; PS00190; CYTOCHROME C; 1.
 KW Electron transport; Photosynthesis; Heme; Signal; 3D-structure;
 KW Complete proteome.
 FT SIGNAL 1 25
 FT CHAIN 26 112 CYTOCHROME C6.
 FT BINDING 39 39 HEME (COVALENT).
 FT BINDING 42 42 HEME (COVALENT).
 FT METAL 43 43 IRON (HEME AXIAL LIGAND).
 FT METAL 83 83 IRON (HEME AXIAL LIGAND).
 FT HELIX 28 35
 FT TURN 36 38
 FT HELIX 39 42
 FT TURN 44 46
 FT TURN 59 63
 FT TURN 69 76
 FT HELIX 80 80
 FT STRAND 81 82
 FT TURN 83 83
 FT STRAND 87 89
 FT HELIX 93 105
 FT TURN 106 107
 FT TURN 110 112
 SQ SEQUENCE 112 AA; 11762 MW; F9C2A62B88CB0305 CRC64;

Query Match 56.9%; Score 37; DB 1; Length 112;
 Best Local Similarity 54.5%; Pred. No. 5.7;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDE 11
 Db 83 MPAFAGRLTDE 93

RESULT 13
 ID CYC6 SYNLU STANDARD; PRT; 112 AA.
 AC Q9FIL9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cytochrome C6 precursor (Soluble cytochrome f) (Cytochrome c553)
 DE (Cytochrome c-553).
 DE PETJ.
 OS Synecococcus vulcanus (Thermosynechococcus vulcanus).
 OC Bacteria; Cyanobacteria; Chroococcales; Thermosynechococcus.
 OX NCBI_TaxID=32053;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Katoh H., Itoh S., Shen J.-R., Ikeuchi M.;
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Functions as an electron carrier between membrane-bound
 CC cytochrome b6f and photosystem I in oxygenic photosynthesis (By
 CC similarity).
 CC -1- SUBUNIT: Monomer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Potential).
 CC -1- PTM: Binds 1 heme group per molecule.
 CC -1- SIMILARITY: Belongs to the cytochrome c family. PetJ subfamily.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AB052598; BAB20065.1; -;
DR HSSP; P56534; IC68.
DR HAMAP; MF 00594; -; 1.
DR InterPro; IPR003088; Cyt_C1.
DR InterPro; IPR008168; Cyt_C1C.
DR InterPro; IPR000345; Cyt_C1C.
DR Pfam; PF00034; cytochrome_c; 1.
DR PRINTS; PR00605; CYTOCHROME_C; 1.
DR PROSITE; PS00190; CYTOCHROME_C; 1.
KW Electron transport; Photosynthesis; Heme; Signal.
FT SIGNAL 1 25 BY SIMILARITY.
FT CHAIN 26 112 CYTOCHROME C6.
FT BINDING 39 39 HEME (COVALENT) (BY SIMILARITY).
FT BINDING 42 42 HEME (COVALENT) (BY SIMILARITY).
FT METAL 43 43 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
FT METAL 83 83 IRON (HEME AXIAL LIGAND) (BY SIMILARITY).
SQ SEQUENCE 112 AA; 11796 MW; F7C24820682503E5 CRC64;

Query Match 56.9%; Score 37; DB 1; Length 112;
Best Local Similarity 54.5%; Pred. No. 5.7;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDE 11
DB 83 MPFAFAGRLTDE 93

RESULT 14

KDSA_NEIMA
ID KDSA_NEIMA STANDARD; PRT; 280 AA.
AC Q9JU48;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 2-dehydro-3-deoxyphosphoactonate aldolase (EC 2.5.1.55) (Phospho-2-
dehydro-3-deoxyoctonate aldolase) (3-deoxy-D-manno-octulosonic acid 8-
phosphate synthetase) (KDO-8-phosphate synthetase) (KDO 8-P synthase)
(KDOPS).
GN KDSA OR NM1493.
OS Neisseria meningitidis (serogroup A).
OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
OC Neisseriaceae; Neisseria.
OX NCBI_TaxID=65699;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=42491 / Serogroup A / Serotype 4A;
RX MEDLINE=20222556; PubMed=10761919;
RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,
Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,
Whitehead S., Spratt B.G., Barrell B.G.;
RT "Complete DNA sequence of a serogroup A strain of Neisseria
meningitidis Z2491".
RL Nature 404:502-506(2000).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + D-arabinose 5-phosphate
+ H(2)O = 2-dehydro-3-deoxy-D-octonate 8-phosphate + phosphate.
CC -1- PATHWAY: Lipopolysaccharide biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the kdsA family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AL162756; CAB84726.1; -;
DR PIR; G81840; G81840.
DR HAMAP; MF 00056; -; 1.
DR InterPro; IPR006218; DAHP1/KDSA.
DR InterPro; IPR006269; KDSA.
DR Pfam; PF00793; DAHP synth.1; 1.
DR TIGRFAMs; TIGR01362; KDO8P synth.1.
KW Lipopolysaccharide biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 280 AA; 30536 MW; 19A4C89A072E4DB0 CRC64;

Query Match 56.9%; Score 37; DB 1; Length 280;
Best Local Similarity 53.8%; Pred. No. 14;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
DB 111 LPAFLARQTDLVA 123

RESULT 15

KDSA_PASMU
ID KDSA_PASMU STANDARD; PRT; 284 AA.
AC P57853;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE 2-dehydro-3-deoxyphosphoactonate aldolase (EC 2.5.1.55) (Phospho-2-
dehydro-3-deoxyoctonate aldolase) (3-deoxy-D-manno-octulosonic acid 8-
phosphate synthetase) (KDO-8-phosphate synthetase) (KDO 8-P synthase)
(KDOPS).
GN KDSA OR PM0558.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida PM70".
RL Proc Natl Acad Sci U S A. 98:3460-3465(2001).
CC -1- CATALYTIC ACTIVITY: Phosphoenolpyruvate + D-arabinose 5-phosphate
+ H(2)O = 2-dehydro-3-deoxy-D-octonate 8-phosphate + phosphate.
CC -1- PATHWAY: Lipopolysaccharide biosynthesis.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: Belongs to the kdsA family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AE006091; AAK02642.1; -;
DR HAMAP; MF 00056; -; 1.
DR InterPro; IPR006218; DAHP1/KDSA.
DR InterPro; IPR006269; KDSA.
DR Pfam; PF00793; DAHP synth.1; 1.
DR TIGRFAMs; TIGR01362; KDO8P synth.1.
KW Lipopolysaccharide biosynthesis; Transferase; Complete proteome.
SQ SEQUENCE 284 AA; 30949 MW; 924696DD1F8FEA01 CRC64;

Query Match 56.9%; Score 37; DB 1; Length 284;
Best Local Similarity 53.8%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13

Db ||: || : : || |
114 LPAFLARQTDLVA 126

Search completed: September 9, 2004, 10:53:51
Job time : 4.69032 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:44:32 ; Search time 6.62581 Seconds
(without alignments)
188.730 Million cell updates/sec

Title: US-10-067-618-6

Perfect score: 65

Sequence: 1 LPSFLGKRTDEAA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	65	100.0	101	2	A48219	calvasculin - huma
2	65	100.0	101	2	S06207	calvasculin - mous
3	62	95.4	101	2	S01759	calvasculin - rat
4	61	93.8	100	2	A53217	placental calcium-
5	55	84.6	55	2	S20330	calvasculin - bovi
6	44	67.7	221	2	S41387	group 3 Lea protei
7	44	67.7	521	2	T05684	hypothetical prote
8	42	64.6	97	2	A30129	S-100 protein, lun
9	42	64.6	98	2	A41988	S-100 calcium-bind
10	40	61.5	423	2	A49772	cuticle protein cu
11	39	60.0	1215	2	H84513	probable disease r
12	38	58.5	234	2	AE0454	endonuclease V (EC
13	38	58.5	644	2	T08340	probable arsenical
14	38	58.5	2109	1	ZLVN	genome polyprotein
15	37	56.9	87	1	CCYC6L	cytochrome c6 - Sy
16	37	56.9	161	2	S71768	low molecularweig
17	37	56.9	161	2	S71769	low molecularweig
18	37	56.9	280	2	G81840	2-dehydro-3-deoxy-
19	37	56.9	384	2	A84955	cell division prot
20	37	56.9	431	2	C87436	major facilitator
21	37	56.9	500	2	A99999	probable portal pr
22	37	56.9	500	2	H85818	hypothetical prote
23	37	56.9	986	2	AB2209	two-component sens
24	36	55.4	84	2	T10348	hypothetical prote
25	36	55.4	171	2	F75556	probable acetyltra
26	36	55.4	205	2	S05015	beta-crystallin B2
27	36	55.4	208	2	C81800	phosphoribosylglyc
28	36	55.4	208	2	B81068	phosphoribosylglyc
29	36	55.4	234	2	G83122	probable permealase

ALIGNMENTS

RESULT 1

A48219

calvasculin - human

N/Alternate names: calcium-binding protein L (CAPL); metastasin; placental calcium-bindin

C/Species: Homo sapiens (man)

C/Date: 02-Jun-1994 #sequence revision 02-Jun-1994 #text_change 13-Aug-1999

C/Accession: A48219; B44470; F01223

R/Engelkamp, D.; Schaefer, B.W.; Mattei, M.G.; Erne, P.; Heizmann, C.W.

Proc. Natl. Acad. Sci. U.S.A. 90, 6547-6551, 1993

A/Title: Six S100 genes are clustered on human chromosome 1q21: identification of two ge

A/Reference number: A48219; MUID:93342029; PMID:8341667

A/Accession: A48219

A/Status: nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 1-101 <ENG>

A/Cross-references: GB:Z18950; NID:g396708; PIDN:CAA79474.1; PID:g396710

R/Engelkamp, D.; Schaefer, B.W.; Erne, P.; Heizmann, C.W.

Biochemistry 31, 10258-10264, 1992

A/Title: S100 alpha, CAPL, and CACY: molecular cloning and expression analysis of three

A/Reference number: A44470; MUID:9304710; PMID:1384693

A/Accession: B44470

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-101 <EN2>

A/Cross-references: GB:M80563; GB:M77499; NID:g179916; PIDN:AAAS1920.1; PID:g179917

A/Experimental source: heart

A/Note: sequence extracted from NCBI backbone (NCBIP:116496)

R/Tomida, Y.; Terasawa, M.; Kobayashi, R.; Hidaka, H.

Biochem. Biophys. Res. Commun. 189, 1310-1316, 1992

A/Title: Calcyclin and calvasculin exist in human platelets.

A/Reference number: PC1222; MUID:93129189; PMID:1482346

A/Accession: PC1223

A/Molecule type: protein

A/Residues: 8-18;36-57 <TOM>

A/Experimental source: platelets

C/Genetics:

A/Gene: GDB:S100A4; CAPL

A/Cross-references: GDB:119748; OMIM:114210

A/Map position: 1q21-1q21

A/Introns: 47/3

C/Superfamily: S-100 protein; calmodulin repeat homology

C/Keywords: calcium binding; BF hand

F/7-41/Domain: calmodulin repeat homology <EF1>

F/50-82/Domain: calmodulin repeat homology <EF2>

Query Match

Best Local Similarity 100.0%; Score 65; DB 2; Length 101;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13

42 LPSFLGKRTDEAA 54

Db

RESULT 2

S06207

calvasculin - mouse

N:Alternate names: calcium-binding protein mts1; calcium-binding protein pEL98; placenta

C:Species: Mus musculus (house mouse)

C>Date: 31-Dec-1990 #sequence revision 31-Dec-1990 #text_change 20-Jun-2000

C:Accession: S06207; JH0097; S07981; A26803; A41411; I48674

R:Ebraldize, A.; Tulchinsky, E.; Grigorian, M.; Afanasyeva, A.; Senin, V.; Revazova, E.;

Genes Dev. 3, 1086-1093, 1989

A:Title: Isolation and characterization of a gene specifically expressed in different me

A:Reference number: S06207; MUID:89378739; PMID:2550322

A:Accession: S06207

A:Molecule type: mRNA

A:Residues: 1-101 <ERR>

A:Cross-references: EMBL:X16190; NID:G54926; PIDN:CAA34316.1; PID:G54927

R:Tulchinsky, E.M.; Grigorian, M.S.; Ebraldize, A.K.; Milshina, N.I.; Lukanidin, E.M.

Gene 87, 219-223, 1990

A:Title: Structure of gene mts1, transcribed in metastatic mouse tumor cells.

A:Reference number: JH0097; MUID:90236313; PMID:2332170

A:Accession: JH0097

A:Molecule type: DNA

A:Residues: 1-101 <TUL>

A:Cross-references: GB:M36578; GB:M36579

A:Experimental source: liver

R:Tulchinsky, B.

submitted to the EMBL Data Library, August 1989

A:Reference number: S07981

A:Accession: S07981

A:Molecule type: DNA

A:Residues: 1-47, 'VSGSXENG', '56-57, 'RTDEAA', <TUL>

A:Cross-references: EMBL:X16094; NID:G53249; PIDN:CAA34224.1; PID:G53250

R:Jackson-Grusby, L.L.; Swiergiel, J.; Linzer, D.I.H.

Nucleic Acids Res. 15, 6677-6690, 1987

A:Title: A growth-related mRNA in cultured mouse cells encodes a placental calcium bindi

A:Reference number: A26803; MUID:87316927; PMID:3628004

A:Accession: A26803

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-101 <JAC>

A:Cross-references: GB:G00310; PIDN:CAA29282.1; PID:G50311

R:Goto, K.; Endo, H.; Fujiyoshi, T.

J. Biochem. 103, 48-53, 1988

A:Title: Cloning of the sequences expressed abundantly in established cell lines: ident

A:Reference number: A41411; MUID:88198109; PMID:3162911

A:Accession: A41411

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-101 <GOT>

A:Cross-references: GB:D00208; NID:G220569; PIDN:BAA00148.1; PID:G220570

R:Tulchinsky, E.; Kramerov, D.; Ford, H.L.; Reshetnyak, E.; Lukanidin, E.; Zain, S.

Oncogene 8, 79-86, 1993

A:Title: Characterization of a positive regulatory element in the mts1 gene.

A:Reference number: I48674; MUID:93141279; PMID:8423998

A:Accession: I48674

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-47, 'VSGSXENG', '48-54 <RES>

A:Cross-references: EMBL:X16094; NID:G53249; PIDN:CAA34224.1; PID:G53250

C:Comment: Gene mts1 is expressed in metastatic cells.

C:Genetics:

A:Introns: 47/3

C:Superfamily: S-100 protein; calmodulin repeat homology

C:Keywords: calcium binding; cancer; EF hand

F:7-41/Domain: calmodulin repeat homology <EF1>

F:50-82/Domain: calmodulin repeat homology <EF2>

Query Match 100.0%; Score 65; DB 2; Length 101;

Best Local Similarity 100.0%; Pred. No. 9.4e-05;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPSEFLGKRTDEAA 13

Db 42 LPSEFLGKRTDEAA 54

RESULT 3

S01759

calvasculin - rat

N:Alternate names: calcium-binding protein p9Ka/42A; gene p9Ka protein

C:Species: Rattus norvegicus (Norway rat)

C>Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text_change 13-Aug-1999

C:Accession: S01759; B31373; S26496; S26497

R:Barraclough, R.; Savin, J.; Dube, S.K.; Rudland, P.S.

J. Mol. Biol. 198, 13-20, 1987

A:Title: Molecular cloning and sequence of the gene for p9Ka : a cultured myoepithelial

A:Reference number: S01759; MUID:88118907; PMID:3430604

A:Accession: S01759

A:Molecule type: DNA

A:Residues: 1-101 <BAR>

A:Cross-references: EMBL:X06916; NID:G56832; PIDN:CAA30014.1; PID:G56833

R:Maslakowski, P.; Shooter, E.M.

Proc. Natl. Acad. Sci. U.S.A. 85, 1277-1281, 1988

A:Title: Nerve growth factor induces the genes for two proteins related to a family of

A:Reference number: A94189; MUID:88125019; PMID:3422491

A:Accession: B31373

A:Molecule type: mRNA

A:Residues: 1-101 <MAS>

A:Cross-references: GB:J03628; NID:G206829; PIDN:AAA42098.1; PID:G206830

R:de Vouge, M.W.; Mukherjee, B.B.

Oncogene 7, 109-119, 1992

A:Title: Transformation of normal rat kidney cells by v-K-ras enhances expression of tra

A:Reference number: S26496; MUID:92158347; PMID:1741158

A:Accession: S26496

A>Status: preliminary; translation not shown

A:Molecule type: mRNA

A:Residues: 65-100 <DEV>

A:Cross-references: EMBL:X64023

A:Accession: S26497

A>Status: preliminary; translation not shown

A:Molecule type: mRNA

A:Residues: 3-97 <DE2>

A:Cross-references: EMBL:X64022

C:Genetics:

A:Gene: p9Ka

A:Introns: 47/3

C:Superfamily: S-100 protein; calmodulin repeat homology

C:Keywords: calcium binding; EF hand

F:50-82/Domain: calmodulin repeat homology <EF2>

Query Match 95.4%; Score 62; DB 2; Length 101;

Best Local Similarity 92.3%; Pred. No. 0.00034;

Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPSEFLGKRTDEAA 13

Db 42 LPSEFLGKRTDEAA 54

RESULT 4

A53217

Placental calcium-binding protein homolog - bovine

C:Species: Bos primigenius taurus (cattle)

C>Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 17-Mar-1999

C:Accession: A53217

R:Polans, A.S.; Palczewski, K.; Asson-Batres, M.A.; Ohguro, H.; Witkowska, D.; Haley, T.

J. Biol. Chem. 269, 6233-6240, 1994

A:Title: Purification and primary structure of capl, an S-100-related calcium-binding

A:Reference number: A53217; MUID:94164991; PMID:8119967

A:Accession: A53217

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-100 <POL>

C:Superfamily: S-100 protein; calmodulin repeat homology
 C:Keywords: acetylated amino end; calcium binding; EF hand
 F:6-40/Domain: calmodulin repeat homology <EF1>
 F:1/Modified site: acetylated amino end (Ala) #status experimental

Query Match 93.8%; Score 61; DB 2; Length 100;
 Best Local Similarity 92.3%; Pred. No. 0.00051;
 Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LPSFLGKRTDEAA 13
 |||||
 Db 41 LPSFLGKRTDETA 53

RESULT 5

S20330
 calvasculin - bovine (fragments)
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 19-Mar-1997 #sequence_revision 29-May-1998 #text_change 29-Sep-1999
 C:Accession: S20330
 R:Watanabe, Y.; Kobayashi, R.; Ishikawa, T.; Hidaka, H.
 Arch. Biochem. Biophys. 292, 563-569, 1992
 A:Title: Isolation and characterization of a calcium-binding protein derived from mRNA
 A:Reference number: S20330; MUID:92117670; PMID:1731618
 A:Accession: S20330
 A:Molecule type: protein
 A:Residues: 1-10; 11-55 <WAT>
 A:Experimental source: aorta
 C:Superfamily: S-100 protein; calmodulin repeat homology
 C:Keywords: calcium binding; dimer; EF hand
 F:38-49/Domain: calcium binding #status predicted <CA2>
 F:39-41,43,45,50/Binding site: calcium (Asp, Asn, Asp, Glu, Glu) #status predicted

Query Match 84.6%; Score 55; DB 2; Length 55;
 Best Local Similarity 84.6%; Pred. No. 0.0037;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LPSFLGKRTDEAA 13
 |||||
 Db 18 LPSFLGKRTDETA 30

RESULT 6

S41387
 group 3 Lea protein MGL3 - maize
 C:Species: Zea mays (maize)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 02-Sep-2000
 C:Accession: S41387
 R:White, C.N.; Rivin, C.J.
 submitted to the EMBL Data Library, January 1994

A:Description: Sequence of the message encoding a Lea Group 3 of maize.
 A:Reference number: S41387
 A:Accession: S41387
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-221 <WHI>

A:Cross-references: EMBL:Z29512; NID:G444044; PIDN:CAA82632.1; PID:el192700; PID:G265437
 C:Superfamily: barley hypothetical protein (clone ES2A)

Query Match 67.7%; Score 44; DB 2; Length 221;
 Best Local Similarity 80.0%; Pred. No. 1.6;
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 SFLGKRTDEA 12
 :|||
 Db 68 SYLGKRTDEA 77

RESULT 7

T05684
 hypothetical protein F20M13.120 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999

C:Accession: T05684

R:Bevan, M.; Wedler, H.; Kutzner, M.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.
 submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15420

A:Accession: T05684

A:Molecule type: DNA

A:Residues: 1-521 <BEV>

A:Cross-references: EMBL:AL035540

A:Experimental source: cultivar Columbia; BAC clone F20M13

C:Genetics:

A:Map position: 4

A:Introns: 41/1; 132/3; 281/3

A>Note: F20M13.120

Query Match 67.7%; Score 44; DB 2; Length 521;
 Best Local Similarity 66.7%; Pred. No. 3.7;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LPSFLGKRTDEA 12
 |||||
 Db 67 LPSYVGKRMDES 78

RESULT 8

A30129
 S-100 protein, lung - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 01-Sep-1995
 C:Accession: A30129
 R:Glennay Jr., J.R.; Kindy, M.S.; Zokas, L.
 J. Cell Biol. 108, 569-578, 1989
 A:Title: Isolation of a new member of the S100 protein family: amino acid sequence, tiss
 A:Reference number: A30129; MUID:89139574; PMID:2521861
 A:Accession: A30129
 A>Status: not compared with conceptual translation
 A:Molecule type: DNA
 A:Residues: 1-97 <GLE>

C:Superfamily: S-100 protein; calmodulin repeat homology
 C:Keywords: calcium binding; EF hand
 F:50-82/Domain: calmodulin repeat homology <EF2>

Query Match 64.6%; Score 42; DB 2; Length 97;
 Best Local Similarity 63.6%; Pred. No. 1.7;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LPSFLGKRTDE 11
 |||||
 Db 42 LPSFVGKVD 52

RESULT 9

A41988
 S-100 calcium-binding protein A2 - human
 N:Alternate names: calcium-binding protein Can19; S-100 calcium-binding protein L
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 29-Aug-1997
 C:Accession: A41988
 R:Lee, S.W.; Tomasetto, C.; Swisshelm, K.; Keyomarsi, K.; Sager, R.
 Proc. Natl. Acad. Sci. U.S.A. 89, 2504-2508, 1992

A:Title: Down-regulation of a member of the S100 gene family in mammary carcinoma cells

A:Reference number: A41988; MUID:92196147; PMID:1372446

A:Accession: A41988

A:Molecule type: mRNA

A:Residues: 1-98 <LEE>

A:Cross-references: GB:M87068; NID:G179896

A:Experimental source: mammary carcinoma cells

A>Note: sequence extracted from NCBI backbone (NCBI:88561, NCBI:88562)

C:Genetics:

A:Gene: GDB:S100A2; CAN19; S100L

A:Cross-references: GDB:211118; OMIM:176993

A:Map position: 1q21-1q21

C:Superfamily: S-100 protein; calmodulin repeat homology

C:Keywords: calcium binding; EF hand

Thu Sep 9 12:03:25 2004

F:8-42/Domain: calmodulin repeat homology <EF1>
 F:51-83/Domain: calmodulin repeat homology <EF2>

Query Match 64.6%; Score 42; DB 2; Length 98;
 Best Local Similarity 63.6%; Pred. No. 1.7;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDE 11
 |||||:
 Db 43 LPSFVGKVD 53

RESULT 10

cuticle protein cut-1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 01-Dec-2000

C:Accession: A49772; S27799

R:Sebastiano, M.; Lassandro, F.; Bazzicalupo, P.

Dev. Biol. 146, 519-530, 1991

A:Title: cut-1 a Caenorhabditis elegans gene coding for a dauer-specific noncollagenous

A:Reference number: A49772; MUID:91323673; PMID:1864469

A:Accession: A49772

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-423 <SEB>

A:Cross-references: GB:M55997

R:Sebastiano, M.; Lassandro, F.; Bazzicalupo, P.

submitted to the EMBL Data Library, July 1991

A:Description: CUT-1 a Caenorhabditis elegans gene coding for a dauer specific non colla

A:Reference number: S27799

A:Accession: S27799

A:Molecule type: DNA

A:Residues: 1-77-78, 'YP', 121-423 <SE2>

A:Cross-references: EMBL:M55997; NID:9156271; PID:9156272

C:Genetics:

A:Gene: CUT-1

A:Introns: 360/3

C:Superfamily: Caenorhabditis elegans hypothetical protein F22B5.3

Query Match 61.5%; Score 40; DB 2; Length 423;
 Best Local Similarity 58.9%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPSFLGKRT 9
 |||||:
 Db 394 LPSFLGKRT 402

RESULT 11

H84513

Probable disease resistance protein [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: H84513

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.

euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: H84513

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1215 <STO>

A:Cross-references: GB:AE002093; NID:96598711; PIDN:AD25848.2; GSPDB:GN00139

C:Genetics:

A:Gene: At2g14080

A:Map position: 2

Query Match 60.0%; Score 39; DB 2; Length 1215;
 Best Local Similarity 70.0%; Pred. No. 73;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPSFLGKRTD 10
 |||||:
 Db 857 LPSFIGKATN 866

RESULT 12

AE0454

endonuclease V (EC 3.1.1.-) [imported] - Yersinia pestis (strain CO92)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Nov-2001

C:Accession: AE0454

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.

deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.

11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AE0454

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-234 <KUR>

A:Cross-references: GB:AL590842; PIDN:CAC93201.1; PID:915981649; GSPDB:GN00175

C:Genetics:

A:Gene: nfi

C:Superfamily: conserved hypothetical protein AF0129

Query Match 58.5%; Score 38; DB 2; Length 234;
 Best Local Similarity 77.8%; Pred. No. 22;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 PSFLGKRTD 10
 |||||:
 Db 224 PSFLGKRTD 232

RESULT 13

T08340

Probable arsenical pump-driving ATPase (EC 3.6.1.-) H1450 - Halobacterium sp. (strain N

C:Species: Halobacterium sp.

A:Variety: strain NRC-1

C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 31-Jan-2000

C:Accession: T08340

R:Ng, W.V.; Ciuffo, S.A.; Smith, T.M.; Bumgarner, R.E.; Baskin, D.; Faust, J.; Hall, B.;

Genome Res. 8, 1131-1141, 1998

A:Title: Snapshot of a large dynamic replicon in a halophilic Archaeon: megaplasmid or

A:Reference number: Z16408; MUID:99063795; PMID:9847077

A:Accession: T08340

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-644 <DAS>

A:Cross-references: EMBL:AF016485; NID:G2822278; PID:G2822401; HALOSP:H1450

A:Experimental source: strain NRC-1

C:Genetics:

A:Gene: arsA; HALOSP:H1450

A:Genome: plasmid pNRC100

C:Function:

A:Description: catalyzes the extrusion of the oxyanions arsenite, antimonite, and arsen

C:Superfamily: Escherichia coli plasmid R773 arsenical pump-driving ATPase

C:Keywords: ATP; hydrolase; P-loop; toxic oxyanion resistance

Query Match 58.5%; Score 38; DB 2; Length 644;
 Best Local Similarity 66.7%; Pred. No. 60;
 Matches 8; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 2 PSFLGKRTDEAA 13
 |||||:
 Db 251 PFFEGKRADEQA 262

RESULT 14

ZLVN

genome polyprotein - vesicular stomatitis Indiana virus (strain Mudd-Summers)

N:Alternate names: L protein
 N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
 C:Species: vesicular stomatitis Indiana virus
 C>Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 29-May-1998
 C:Accession: A04119
 R:Schubert, M.; Harmison, G.G.; Meier, E.
 J. Virol. 51, 505-514, 1984
 A:Title: Primary structure of the vesicular stomatitis virus polymerase (L) gene: evidence
 A:Reference number: A04119; PMID:6086959
 A:Accession: A04119
 A:Molecule type: genomic RNA
 A:Residues: 1-2109 <SCH>
 A:Cross-references: GB:K02378
 C:Genetics:
 A:Gene: L
 C:Superfamily: rhabdovirus L protein
 C:Keywords: nucleotidyltransferase; RNA binding; RNA biosynthesis; transmembrane protein
 F:530-549/Region: RNA binding #status predicted
 F:1959-1978/Domain: transmembrane #status predicted <TMN>

Query Match 58.5%; Score 38; DB 1; Length 2109;
 Best Local Similarity 50.0%; Pred. No. 1.9e+02;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEA 12
 ||:|||:|:
 Db 1149 LPAVLGSKTSES 1160

RESULT 15
 CCYC6L
 cytochrome c6 - Synecococcus lividus
 N:Alternate names: cytochrome c553; soluble cytochrome f
 C:Species: Synecococcus lividus
 C>Date: 31-Jan-1980 #sequence_revision 31-Jan-1980 #text_change 03-Mar-2000
 C:Accession: A00106
 R:Borden, D.; Margoliash, E.
 submitted to the Atlas, December 1979
 A:Reference number: A00106
 A:Accession: A00106
 A:Molecule type: protein
 A:Residues: 1-87 <BOR>
 C:Comment: Cytochrome c6 substitutes for plastocyanin in copper-deficient blue-green alga
 C:Superfamily: cytochrome c6; cytochrome c6 homology
 C:Keywords: chromoprotein; electron transfer; heme; iron; metalloprotein; photosynthesis
 F:4-77/Domain: cytochrome c6 homology <CYC>
 F:14,17/Binding site: heme (Cys) (covalent) #status predicted
 F:18,58/Binding site: heme iron (His, Met) (axial ligands) #status predicted

Query Match 56.9%; Score 37; DB 1; Length 87;
 Best Local Similarity 54.5%; Pred. No. 13;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDE 11
 :|||:|:
 Db 58 MPAPAGRLTDE 68

Search completed: September 9, 2004, 10:59:41
 Job time : 8.62581 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:47:43 ; search time 7.29677 Seconds
(without alignments)
91.977 Million cell updates/sec

Title: US-10-067-618-6

Perfect score: 65

Sequence: 1 LPSFLGKRTDEAA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA.*
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCUTUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles!.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	65	100.0	13	1	US-08-190-560-6
2	65	100.0	13	1	US-08-469-277-6
3	65	100.0	13	2	US-08-468-946-6
4	65	100.0	13	2	US-08-468-942-6
5	65	100.0	13	4	US-09-298-625-6
6	65	100.0	101	1	US-08-190-560-2
7	65	100.0	101	1	US-08-469-277-2
8	65	100.0	101	2	US-08-468-946-2
9	65	100.0	101	2	US-08-468-942-2
10	65	100.0	101	4	US-09-298-625-2
11	42	64.6	97	1	US-07-662-198B-2
12	38	58.5	2109	3	US-08-646-695-6
13	38	58.5	2109	5	PCT-US96-06053-6
14	37	56.9	832	4	US-09-252-991A-31601
15	36	55.4	199	4	US-09-489-039A-8835
16	36	55.4	285	4	US-09-328-352-7570
17	36	55.4	287	4	US-09-489-039A-10380
18	36	55.4	290	4	US-09-540-236-2355
19	36	55.4	292	4	US-09-107-532A-5310
20	36	55.4	302	4	US-09-252-991A-21019
21	36	55.4	303	4	US-09-543-681A-4446
22	36	55.4	317	4	US-09-252-991A-30686
23	36	55.4	450	4	US-09-252-991A-20484
24	36	55.4	457	3	US-09-416-213-2
25	36	55.4	457	4	US-09-416-214-2
26	36	55.4	457	4	US-09-035-676-2
27	36	55.4	513	4	US-09-489-039A-9121

28	36	55.4	550	4	US-09-786-261-2	Sequence 2, Appli
29	36	55.4	841	4	US-09-252-991A-26919	Sequence 28919, A
30	36	55.4	1494	3	US-08-755-587-186	Sequence 186, App
31	35	53.8	182	4	US-09-252-991A-17938	Sequence 17938, A
32	35	53.8	218	4	US-09-134-000C-4489	Sequence 4489, Ap
33	35	53.8	270	4	US-09-134-000C-5024	Sequence 5024, Ap
34	35	53.8	391	4	US-09-252-991A-28342	Sequence 28342, A
35	35	53.8	402	2	US-08-403-852D-19	Sequence 19, Appl
36	35	53.8	402	3	US-08-510-648B-20	Sequence 19, Appl
37	35	53.8	402	3	US-09-231-818-19	Sequence 20, Appl
38	35	53.8	402	4	US-09-635-359B-19	Sequence 19, Appl
39	35	53.8	1580	2	US-08-804-227C-11	Sequence 11, Appl
40	35	53.8	1580	2	US-08-804-198-5	Sequence 5, Appli
41	34	52.3	96	4	US-09-370-838-77	Sequence 77, Appl
42	34	52.3	332	4	US-09-252-991A-28896	Sequence 28896, A
43	34	52.3	364	4	US-09-328-352-4659	Sequence 4659, Ap
44	34	52.3	417	4	US-09-252-991A-24107	Sequence 24107, A
45	34	52.3	510	4	US-09-543-681A-6077	Sequence 6077, Ap

ALIGNMENTS

RESULT 1
US-08-190-560-6
; Sequence 6, Application US/08190560
; Patent No. 5798257
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE WTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 13 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-190-560-6

Query Match 100.0%; Score 65; DB 1; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.9e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13

Db 1 LPSFLGKRTDEAA 13

us-10-067-618-6.ra1


```

; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-190-560-2

Query Match 100.0%; Score 65; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels

QY 1 LPSFLGKRTDEAA 13
DB 42 LPSFLGKRTDEAA 54
|||||
|||||

RESULT 7
US-08-469-277-2
; Sequence 2, Application US/08469277
; Patent No. 5801142
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukandin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,277
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366

```

us-10-067-618-6.ra1

Thu Sep 9 12:03:24 2004

TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 101 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-469-277-2

Query Match 100.0%; Score 65; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPSFLGKRTDEAA 13
Db 42 LPSFLGKRTDEAA 54

RESULT 8
US-08-468-946-2
Sequence 2, Application US/08468946
Patent No. 5843686
GENERAL INFORMATION:
APPLICANT: Zain, Sayeeda
TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
TITLE OF INVENTION: THE MTS-1 GENE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,946
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,560
FILING DATE: 31-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 78792Y
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 101 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match 100.0%; Score 65; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPSFLGKRTDEAA 13
Db 42 LPSFLGKRTDEAA 54

US-08-468-946-2
Query Match 100.0%; Score 65; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPSFLGKRTDEAA 13
Db 42 LPSFLGKRTDEAA 54

RESULT 9
US-08-468-942-2
Sequence 2, Application US/08468942
Patent No. 5965360
GENERAL INFORMATION:
APPLICANT: Zain, Sayeeda
TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
TITLE OF INVENTION: THE MTS-1 GENE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,942
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,560
FILING DATE: 31-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 78792Y
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 101 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match 100.0%; Score 65; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.7e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPSFLGKRTDEAA 13
Db 42 LPSFLGKRTDEAA 54

RESULT 10
US-09-238-625-2
Sequence 2, Application US/09298625
Patent No. 6638504
GENERAL INFORMATION:
APPLICANT: Lukanidin, Eugene
TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
FILE REFERENCE: 78792VAIII-Z
CURRENT APPLICATION NUMBER: US/09/298,625
CURRENT FILING DATE: 1999-04-23
PRIOR APPLICATION NUMBER: 08/468,942
PRIOR FILING DATE: 1995-06-06
PRIOR APPLICATION NUMBER: 08/190,560
PRIOR FILING DATE: 1994-01-31
PRIOR APPLICATION NUMBER: 07/981,455
PRIOR FILING DATE: 1992-11-25
PRIOR APPLICATION NUMBER: 07/550,600

; PRIOR FILING DATE: 1990-07-09
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 101
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-298-625-2

Query Match 100.0%; Score 65; DB 4; Length 101;
 Best Local Similarity 100.0%; Pred. No. 2.7e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
 |||||
 DB 42 LPSFLGKRTDEAA 54

RESULT 11
 US-07-662-198B-2
 ; Sequence 2, Application US/07662198B
 ; Patent No. 5262528
 ; GENERAL INFORMATION:
 ; APPLICANT: Sager, Ruth
 ; APPLICANT: Lee, Sam W.
 ; APPLICANT: Tomasetto, Catherine
 ; TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: Massachusetts
 ; COUNTRY: U.S.A.
 ; ZIP: 02110-2804

; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 ; COMPUTER: IBM PS/2 Model 50Z or 55SX
 ; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
 ; SOFTWARE: WordPerfect (Version 5.0)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/662,198B
 ; FILING DATE: 19910228
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Fraser, Janis K.
 ; REGISTRATION NUMBER: 34,819
 ; REFERENCE/DOCKET NUMBER: 00530/049001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (617) 542-5070
 ; TELEFAX: (617) 542-8906
 ; TELEX: 200154

; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 97
 ; TYPE: AMINO ACID
 ; STRANDEDNESS:
 ; TOPOLOGY: linear
 US-07-662-198B-2

Query Match 64.6%; Score 42; DB 1; Length 97;
 Best Local Similarity 63.6%; Pred. No. 0.78;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDE 11
 |||||
 DB 42 LPSFVGKVE 52

RESULT 12

US-08-646-695-6
 ; Sequence 6, Application US/08646695
 ; Patent No. 6168943
 ; GENERAL INFORMATION:
 ; APPLICANT: Rose, John K.
 ; TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR
 ; TITLE OF INVENTION: USES
 ; NUMBER OF SEQUENCES: 44
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PENNIE & EDMONDS
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/646,695
 ; FILING DATE: On Even Date Herewith
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Misrock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 6523-008
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9741/8864
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 2109 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 US-08-646-695-6

Query Match 58.5%; Score 38; DB 3; Length 2109;
 Best Local Similarity 50.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEA 12
 ||::|||
 DB 1149 LPAYLGSKTSES 1160

RESULT 13
 PCT-US96-06053-6
 ; Sequence 6, Application PC/TUS9606053
 ; GENERAL INFORMATION:
 ; APPLICANT: Yale University
 ; TITLE OF INVENTION: RECOMBINANT VESICULOVIRUSES AND THEIR
 ; TITLE OF INVENTION: USES
 ; NUMBER OF SEQUENCES: 41
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PENNIE & EDMONDS
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US96/06053
 ; FILING DATE: 01-MAY-1996
 ; CLASSIFICATION:

us-10-067-618-6.ra1

Thu Sep 9 12:03:24 2004

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6523-009-228
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2109 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-06053-6

Query Match 58.5%; Score 38; DB 5; Length 2109;
Best Local Similarity 50.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEA 12
Db 1149 LPVILGSKTSES 1160

RESULT 14
US-09-252-991A-31601
; Sequence 31601, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31601
; LENGTH: 832
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31601

Query Match 56.9%; Score 37; DB 4; Length 832;
Best Local Similarity 58.3%; Pred. No. 75;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 PSLGKRTDEAA 13
Db 79 PSLGKSRDRTS 90

RESULT 15
US-09-489-039A-8835
; Sequence 8835, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8835
; LENGTH: 199
; TYPE: PRT

```

```

; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8835

Query Match 55.4%; Score 36; DB 4; Length 199;
Best Local Similarity 61.5%; Pred. No. 25;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
Db 137 LPSFLGHWNGDNA 149

Search completed: September 9, 2004, 11:01:12
Job time : 8.29677 secs

```

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:53:18 ; Search time 17.4452 Seconds
(without alignments)
238.975 Million cell updates/sec

Title: US-10-067-618-6
Perfect score: 65
Sequence: 1 LPSFLGKRTDEAA 13

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09E_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	100.0	13	13 US-10-067-618-6	Sequence 6, Appli
2	65	100.0	13	13 US-10-135-152-6	Sequence 6, Appli
3	65	100.0	101	9 US-09-393-433-1	Sequence 1, Appli
4	65	100.0	101	9 US-09-393-433-2	Sequence 2, Appli
5	65	100.0	101	9 US-09-781-509-1	Sequence 1, Appli
6	65	100.0	101	9 US-09-781-509-2	Sequence 2, Appli
7	65	100.0	101	12 US-10-087-192-1158	Sequence 1158, Ap
8	65	100.0	101	13 US-10-067-618-2	Sequence 2, Appli
9	65	100.0	101	13 US-10-135-152-1	Sequence 2, Appli
10	65	100.0	101	14 US-10-269-643-1	Sequence 1, Appli
11	65	100.0	101	14 US-10-269-643-2	Sequence 2, Appli
12	65	100.0	119	12 US-10-087-192-1155	Sequence 1155, Ap
13	42	64.6	97	14 US-10-097-340-274	Sequence 274, App
14	42	64.6	97	14 US-10-171-311-206	Sequence 206, App
15	42	64.6	97	15 US-10-236-031B-40	Sequence 40, Appl

16	40	61.5	195	15	US-10-310-154-399	Sequence 399, App
17	40	61.5	298	12	US-10-425-114-68843	Sequence 68843, A
18	40	61.5	379	15	US-10-369-493-10136	Sequence 10336, A
19	40	61.5	637	15	US-10-369-493-12290	Sequence 12290, A
20	39	60.0	163	16	US-10-437-963-142847	Sequence 142847, A
21	38	58.5	1753	16	US-10-743-649-51	Sequence 51, Appl
22	38	58.5	2109	16	US-10-743-649-50	Sequence 50, Appl
23	38	58.5	2109	16	US-10-743-649-52	Sequence 52, Appl
24	37	56.9	76	12	US-10-424-599-163103	Sequence 163103, A
25	37	56.9	82	16	US-10-437-963-200896	Sequence 200896, A
26	37	56.9	146	9	US-09-893-737-40	Sequence 40, Appl
27	37	56.9	151	14	US-10-017-161-1166	Sequence 1166, Ap
28	37	56.9	151	14	US-10-017-161-1400	Sequence 1400, Ap
29	37	56.9	153	16	US-10-437-963-186463	Sequence 186463, A
30	37	56.9	233	16	US-10-437-963-158815	Sequence 158815, A
31	37	56.9	280	12	US-10-282-122A-65245	Sequence 65245, A
32	37	56.9	280	12	US-10-282-122A-65924	Sequence 65924, A
33	37	56.9	283	12	US-10-282-122A-67026	Sequence 67026, A
34	37	56.9	284	16	US-10-437-963-199693	Sequence 199693, A
35	37	56.9	293	15	US-10-166-653-12	Sequence 12, Appl
36	37	56.9	353	16	US-10-437-963-117164	Sequence 117164, A
37	37	56.9	505	12	US-10-424-599-158246	Sequence 158246, A
38	37	56.9	818	16	US-10-437-963-115610	Sequence 115610, A
39	37	56.9	818	16	US-09-774-639-312	Sequence 312, App
40	36	55.4	24	10	US-09-969-730-239	Sequence 239, App
41	36	55.4	24	10	US-10-621-363-239	Sequence 239, App
42	36	55.4	24	16	US-10-437-963-113438	Sequence 113438, A
43	36	55.4	73	16	US-10-263-828-83	Sequence 83, Appl
44	36	55.4	106	14	US-10-424-599-223762	Sequence 223762, A
45	36	55.4	127	12		

ALIGNMENTS

RESULT 1
US-10-067-618-6
; Sequence 6, Application US/10067618
; Publication No. US20020169110A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 78792YAI11-Z
; CURRENT APPLICATION NUMBER: US/10/067,618
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/298,625
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide
US-10-067-618-6

Query Match 100.0%; Score 65; DB 13; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
|||||
Db 1 LPSFLGKRTDEAA 13

Thu Sep 9 12:03:24 2004

GENERAL INFORMATION:
 ; APPLICANT: Lukanidin, Eugene
 ; APPLICANT: Bock, Elisabeth M.
 ; APPLICANT: Berezen, Vladimir
 ; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
 ; FILE REFERENCE: RCT
 ; CURRENT APPLICATION NUMBER: US/09/393,433
 ; CURRENT FILING DATE: 1999-09-10
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 101
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-393-433-2

Query Match 100.0%; Score 65; DB 9; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.00011;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
 |||||
 Db 42 LPSFLGKRTDEAA 54

RESULT 5
 US-09-781-509-1
 ; Sequence 1, Application US/09781509
 ; Patent No. US20020099010A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lukanidin, Eugene
 ; APPLICANT: Bock, Elisabeth M.
 ; APPLICANT: Berezen, Vladimir
 ; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
 ; FILE REFERENCE: RCT
 ; CURRENT APPLICATION NUMBER: US/09/781,509
 ; CURRENT FILING DATE: 2001-02-12
 ; PRIOR APPLICATION NUMBER: 09/393,433
 ; PRIOR FILING DATE: 1999-09-10
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 101
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-781-509-1

Query Match 100.0%; Score 65; DB 9; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.00011;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
 |||||
 Db 42 LPSFLGKRTDEAA 54

RESULT 6
 US-09-781-509-2
 ; Sequence 2, Application US/09781509
 ; Patent No. US20020099010A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lukanidin, Eugene
 ; APPLICANT: Bock, Elisabeth M.
 ; APPLICANT: Berezen, Vladimir
 ; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
 ; FILE REFERENCE: RCT
 ; CURRENT APPLICATION NUMBER: US/09/781,509
 ; CURRENT FILING DATE: 2001-02-12
 ; PRIOR APPLICATION NUMBER: 09/393,433
 ; PRIOR FILING DATE: 1999-09-10
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 2

US-10-135-152-6
 ; Sequence 6, Application US/10135152
 ; Publication No. US20020172680A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lukanidin, Eugene
 ; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
 ; FILE REFERENCE: 78792VAILI-Z
 ; CURRENT APPLICATION NUMBER: US/10/135,152
 ; CURRENT FILING DATE: 2002-04-29
 ; PRIOR APPLICATION NUMBER: US/09/298,625
 ; PRIOR FILING DATE: 1999-04-23
 ; PRIOR APPLICATION NUMBER: 08/468,942
 ; PRIOR FILING DATE: 1995-06-06
 ; PRIOR APPLICATION NUMBER: 08/190,560
 ; PRIOR FILING DATE: 1994-01-31
 ; PRIOR APPLICATION NUMBER: 07/981,455
 ; PRIOR FILING DATE: 1992-11-25
 ; PRIOR APPLICATION NUMBER: 07/550,600
 ; PRIOR FILING DATE: 1990-07-09
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 13
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:peptide
 ; US-10-135-152-6

Query Match 100.0%; Score 65; DB 13; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.2e-05;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
 |||||
 Db 1 LPSFLGKRTDEAA 13

RESULT 3
 US-09-393-433-1
 ; Sequence 1, Application US/09393433
 ; Patent No. US20010011126A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lukanidin, Eugene
 ; APPLICANT: Bock, Elisabeth M.
 ; APPLICANT: Berezen, Vladimir
 ; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
 ; FILE REFERENCE: RCT
 ; CURRENT APPLICATION NUMBER: US/09/393,433
 ; CURRENT FILING DATE: 1999-09-10
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 101
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-393-433-1

Query Match 100.0%; Score 65; DB 9; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.00011;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
 |||||
 Db 42 LPSFLGKRTDEAA 54

RESULT 4
 US-09-393-433-2
 ; Sequence 2, Application US/09393433
 ; Patent No. US20010011126A1

```
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-781-509-2

Query Match      100.0%; Score 65; DB 9; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
Db 42 LPSFLGKRTDEAA 54

RESULT 7
US-10-087-192-1158
; Sequence 1158, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 7879ZYAIII-Z
; CURRENT APPLICATION NUMBER: US/10/087,192
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1158
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-1158

Query Match      100.0%; Score 65; DB 12; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
Db 42 LPSFLGKRTDEAA 54

RESULT 8
US-10-067-618-2
; Sequence 2, Application US/10067618
; Publication No. US20020169110A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; FILE REFERENCE: 7879ZYAIII-Z
; CURRENT APPLICATION NUMBER: US/10/067,618
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/298,625
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-067-618-2

Query Match      100.0%; Score 65; DB 13; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
Db 42 LPSFLGKRTDEAA 54

RESULT 9
US-10-135-152-2
; Sequence 2, Application US/10135152
; Publication No. US20020172680A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 7879ZYAIII-Z
; CURRENT APPLICATION NUMBER: US/10/135,152
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US/09/298,625
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-135-152-2

Query Match      100.0%; Score 65; DB 13; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
Db 42 LPSFLGKRTDEAA 54

RESULT 10
US-10-269-643-1
; Sequence 1, Application US/10269643
; Publication No. US20030100503A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezin, Vladimir
; TITLE OF INVENTION: TITLES OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/10/269,643
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/781,509
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 09/393,433
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-269-643-1

Query Match      100.0%; Score 65; DB 14; Length 101;
```

Thu Sep 9 12:03:24 2004

Best Local Similarity 100.0%; Pred. No. 0.00011; Indels 0; Gaps 0; Mismatches 0;

Qy 1 LPSFLGKRTDEAA 13
Db 42 LPSFLGKRTDEAA 54

RESULT 11

US-10-269-643-2
; Sequence 2, Application US/10269643
; Publication No. US20030100503A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berzen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/10/269,643
; CURRENT FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/781,509
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 09/393,433
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-269-643-2

Query Match 100.0%; Score 65; DB 14; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPSFLGKRTDEAA 13
Db 42 LPSFLGKRTDEAA 54

RESULT 12

US-10-087-192-1155
; Sequence 1155, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-1155

Query Match 100.0%; Score 65; DB 12; Length 119;
Best Local Similarity 100.0%; Pred. No. 0.00013;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LPSFLGKRTDEAA 13
Db 60 LPSFLGKRTDEAA 72

RESULT 13

US-10-097-340-274
; Sequence 274, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPURU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KANATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,

; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-274

Query Match 64.6%; Score 42; DB 14; Length 97;
Best Local Similarity 63.6%; Pred. No. 3.3;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LPSFLGKRTDB 11
Db 42 LPSFVGKVD 52

RESULT 14

US-10-171-311-206
; Sequence 206, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kanatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoers, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY

; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 206
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-171-311-206

Query Match 64.6%; Score 42; DB 14; Length 97;
Best Local Similarity 63.6%; Pred. No. 3.3;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 LPSFLGKRTDE 11
Db 42 LPSFVGKVD 52

RESULT 15
US-10-236-031B-40
; Sequence 40, Application US/10236031B
; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullans, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B00801/70265 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-236-031B-40

Query Match 64.6%; Score 42; DB 15; Length 97;
Best Local Similarity 63.6%; Pred. No. 3.3;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
QY 1 LPSFLGKRTDE 11
Db 42 LPSFVGKVD 52

Search completed: September 9, 2004, 11:04:46
Job time : 17.4452 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:32:01 ; Search time 27.3419 Seconds
(without alignments)
134.340 Million cell updates/sec

Title: US-10-067-618-6
Perfect score: 65
Sequence: 1 LPSFLGKRTDEAA 13

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	65	100.0	13	2	AAR80456 Human mts
2	65	100.0	13	3	AAB37435 Human mts
3	65	100.0	13	6	ABU08520 Human mts
4	65	100.0	101	2	AAR20560 Human mts
5	65	100.0	101	2	AAR80453 Human mts
6	65	100.0	101	3	AAB45534 Human S10
7	65	100.0	101	3	AAB37432 Human mts
8	65	100.0	101	4	AAB72387 Murine Mt
9	65	100.0	101	4	AAB72386 Human Mts
10	65	100.0	101	6	ABU08513 Human mts
11	65	100.0	101	7	ADD14157 Human src
12	44	67.7	521	5	ABP93293 Herbicida
13	42	64.6	97	2	AAR27058 Sequence
14	42	64.6	97	2	AAR26406 Sequence
15	42	64.6	97	3	AAB45532 Human S10
16	42	64.6	97	5	ABG96408 Human ova
17	42	64.6	97	6	ABR92148 Human cer
18	42	64.6	97	7	ADB70348 S100 calc
19	42	64.6	97	7	ABR92148 Human cer
20	42	64.6	97	7	ABR92148 Human cer
21	42	64.6	97	7	ABR92148 Human cer
22	42	64.6	98	6	ABP71986 Human pro
23	42	64.6	98	6	ABU56415 Lung canc
24	42	64.6	98	6	ABU56415 Lung canc
25	38	58.5	94	4	AAU47526 Propionib

26	38	58.5	94	6	ABM44045
27	38	58.5	1753	4	AAB20291
28	38	58.5	2109	2	AAB20291
29	38	58.5	2109	2	AAB20292
30	38	58.5	2109	4	AAB59299
31	37	57.7	483	4	AAM40768
32	37	56.9	146	5	AAU83094
33	37	56.9	280	6	ABP78839
34	37	56.9	280	6	ABP78839
35	37	56.9	280	6	ABU37321
36	37	56.9	280	6	ABU38000
37	37	56.9	283	6	ABU27534
38	37	56.9	284	6	ABU39102
39	37	56.9	317	4	ABG03344
40	37	56.9	363	4	ABG14644
41	37	56.9	500	7	ADC00532
42	36	55.4	784	5	ABB93947
43	36	55.4	24	7	ADB47956
44	36	55.4	42	7	ADC40606
45	36	55.4	100	4	AAB87624
			128	6	ADA55220

ALIGNMENTS

RESULT 1
AAR80456
ID AAR80456 standard; peptide; 13 AA.

XX AC AAR80456;

XX DT 27-DEC-1995 (first entry)

XX DE Human mts-1 (42-54) peptide 3.

XX KW Metastasis; cancer; mts-1 gene; tumour; therapy; antigen; antibody.

XX OS Synthetic.

XX PN WO9520656-A1.

XX PD 03-AUG-1995.

XX PF 31-JAN-1995; 95WO-US001214.

XX PR 31-JAN-1994; 94US-00190560.

XX (RESE) RESEARCH CORP TECHNOLOGIES INC.

XX PA Zain S, Lukanidin E;

XX PI WPI; 1995-275441/36.

XX DR Nucleic acid encoding human mts-1, antigenic fragments and antibodies - useful for diagnosis of malignant cancer and metastatic potential of tumour cells.

XX PT Disclosure; Page 24; 124pp; English.

XX PS Antigenic peptides 1-4 (given in AAR80454-57) of human mts-1 protein (AAR80453) were used to raise polyclonal and monoclonal antibodies that detect the presence of mts-1 in tissue samples, esp. metastatic cells. Peptides 1, 3 and 4 are unique to mts-1, while peptide 2 generates antibodies reactive with many calcium binding proteins

XX SQ Sequence 13 AA;

Query Match 100.0%; Score 65; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13

|||||

```

Db      1 LPSFLGKRTDEAA 13

RESULT 2
AAB37435
ID AAB37435 standard; peptide; 13 AA.
XX
XX
AC AAB37435;
XX
DT 21-FEB-2001 (first entry)
XX
DE Human mts-1 peptide fragment #3.
XX
XX Human; mts-1; cytostatic; cancer; metastasis.
XX
XX Homo sapiens.
XX
XX WO200064475-A1.
XX
XX 02-NOV-2000.
XX
XX 20-APR-2000; 2000WO-US011006.
XX
XX 23-APR-1999; 99US-00298625.
XX
XX (RESE ) RESEARCH CORP TECHNOLOGIES INC.
XX
XX Lukanidin E;
XX
XX WPI; 2000-687266/67.
XX
XX Treating cancer and preventing metastasis comprises administration of an
XX anti-mts-1 protein antibody or antisense oligonucleotide.
XX
XX Example 9; Page 27; 153pp; English.
XX
XX The present invention relates to methods for treating cancer and
XX preventing metastases, comprising the administration of a composition
XX directed against the mts-1 protein (see AAC68131-C68132 and AAB37432).
XX Mts-1 protein is a calcium-binding protein, and is thought to have a role
XX in myoepithelial cell differentiation. The present sequence is a peptide
XX fragment of the human mts-1 protein. This peptide was used to generate
XX antibodies against mts-1 protein, which can be used to detect mts-1
XX protein in clinical specimens
XX
XX Sequence 13 AA;

Query Match 100.0%; Score 65; DB 3; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
DB 1 LPSFLGKRTDEAA 13

RESULT 3
ABU08520
ID ABU08520 standard; peptide; 13 AA.
XX
XX ABU08520;
XX
XX 27-MAY-2003 (first entry)
XX
XX Human mts-1 protein, antigenic peptide #3.
XX
XX Human; mts-1; metastatic cancer; p53; cytostatic; gene therapy; antigen.
XX
XX Homo sapiens.
XX
XX US2002172680-A1.
XX
XX 21-NOV-2002.
XX

XX
XX 29-APR-2002; 2002US-00135152.
XX
XX 09-JUL-1990; 90US-00550600.
XX
XX 25-NOV-1992; 92US-00981455.
XX
XX 31-JAN-1994; 94US-00190560.
XX
XX 06-JUN-1995; 95US-00468942.
XX
XX 23-APR-1999; 99US-00298635.
XX
XX (LUKA/) LUKANIDIN E.
XX
XX Lukanidin E;
XX
XX WPI; 2003-328422/31.
XX
XX Treating cancer by administering a reagent directed against the mts-1
XX protein or an oligonucleotide capable of binding to mts-1 mRNA.
XX
XX Example 9; Page 19; 64pp; English.
XX
XX The invention relates to treating cancer comprising administering a
XX reagent directed against the mts-1 (not defined) protein or an
XX oligonucleotide capable of binding to mts-1 mRNA. Also include are a
XX method of inactivating, destroying or nullifying a mts-1 protein or cells
XX and a method of inhibiting metastasis in a cancerous cell. Experiments
XX showed that mts-1 inhibited the phosphorylation of full-size p53 and the
XX C-terminal protein fragment by PKC. Addition of the same concentrations
XX of mts-1 to the PKC reaction mixture did not affect the phosphorylation
XX of the N-terminal and DNA-binding domains of p53. The method is useful
XX for treating cancers associated with the mts-1 gene. The present sequence
XX represents a human mts-1 protein antigenic peptide
XX
XX Sequence 13 AA;

Query Match 100.0%; Score 65; DB 6; Length 13;
Best Local Similarity 100.0%; Pred. No. 2.6e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDEAA 13
DB 1 LPSFLGKRTDEAA 13

RESULT 4
AAR20560
ID AAR20560 standard; protein; 101 AA.
XX
XX AAR20560;
XX
XX 25-MAR-2003 (revised)
XX
XX 14-MAY-1992 (first entry)
XX
XX Human mts protein.
XX
XX Metastatic cancer; antibodies; mouse; lung; liver; kidney; thyroid;
XX breast cancer; cell growth.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT Protein 1..101
XX FT /label= mts-1
XX FT Peptide 2..11
XX FT /note= "antigenic"
XX FT Peptide 22..37
XX FT /note= "antigenic; calcium binding domain"
XX FT Peptide 42..54
XX FT /note= "antigenic"
XX FT Peptide 87..101
XX FT /note= "antigenic"
XX
XX WO9200757-A.
XX

```

PD 23-JAN-1992.
 XX
 PF 09-JUL-1990; 90US-00550600.
 XX
 PR 09-JUL-1990; 90US-00550600.
 XX
 PA (RESE) RESEARCH CORP TECHNOLOGIES INC.
 XX
 PI Zain S, Lukanidin E;
 XX
 DR WPI; 1992-056647/07.
 DR N-PSDB; AAQ20506.
 XX
 PT Metastatic cancer diagnosis by detection of mts-1 gene or protein - using
 PT antibody treatment of cancer and tumours of e.g. kidney, thyroid, lung
 PT and liver.
 XX
 PS Claim 9; Fig 2; 82pp; English.
 XX
 CC The sequence was deduced from the DNA sequence obtd. by screening a human
 CC cDNA library with mouse mts-1 cDNA probes. The antigenic Mts-1 peptides
 CC (see features) derived from the protein and anti- bodies raised to them
 CC are useful in the diagnosis of metastatic cancer, e.g. lung, kidney,
 CC thyroid or breast cancer. The peptide comprising the calcium binding
 CC generates antibodies reactive with many members of the calcium binding
 CC protein family; the other three peptides are unique to mts-1 and generate
 CC antibodies specific only for this protein. Cell lines capable of
 CC expressing mts-1 are useful as model systems for in vitro and in vivo
 CC anti-metastasis drug screening. Pharmaceutical compns. contg. the mts-1
 CC protein or anti-cancer reagents may be used to promote cell growth, or
 CC for treating cancer, respectively. (Updated on 25-MAR-2003 to correct PA
 CC field.)
 XX
 SQ Sequence 101 AA;
 Query Match 100.0%; Score 65; DB 2; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.00023;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LPSFLGKRTDEAA 13
 Db |||||
 42 LPSFLGKRTDEAA 54
 RESULT 5
 AAR80453
 ID AAR80453 standard; protein; 101 AA.
 XX
 AC AAR80453;
 XX
 DT 27-DEC-1995 (first entry)
 XX
 DE Human mts-1 protein.
 XX
 KW Metastasis; cancer; mts-1 gene; tumour; therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9520656-A1.
 XX
 PF 03-AUG-1995.
 XX
 PD 31-JAN-1995; 95WO-US001214.
 XX
 PF 31-JAN-1994; 94US-00190560.
 XX
 PR (RESE) RESEARCH CORP TECHNOLOGIES INC.
 XX
 PA Zain S, Lukanidin E;
 XX
 PI WPI; 1995-275441/36.
 XX
 DR N-PSDB; AAQ99177.
 XX

PT Nucleic acid encoding human mts-1, antigenic fragments and antibodies -
 PT useful for diagnosis of malignant cancer and metastatic potential of
 PT tumour cells.
 XX
 PS Claim 12; Page 92; 124pp; English.
 XX
 CC A human cDNA library was constructed in lambda-gt10 using poly(A)+ RNA
 CC prep'd. from HeLa cells. The library was screened with a 32P- labeled
 CC mouse mts-1 vDNA probe. A clone was obtd. which comprised the full-length
 CC human mts-1 gene. The encoded protein is used for the diagnosis or
 CC therapy of cancer, and to raise antibodies
 XX
 SQ Sequence 101 AA;
 Query Match 100.0%; Score 65; DB 2; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.00023;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LPSFLGKRTDEAA 13
 Db |||||
 42 LPSFLGKRTDEAA 54
 RESULT 6
 AAB45534
 ID AAB45534 standard; protein; 101 AA.
 XX
 AC AAB45534;
 XX
 DT 22-FEB-2001 (first entry)
 XX
 DE Human S100A4 protein.
 XX
 KW S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;
 KW calcium-binding protein; calcium homeostasis; cardiac muscle;
 KW pumping capacity; myocardial cell; systolic calcium ion release;
 KW sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;
 KW valve defect.
 XX
 OS Homo sapiens.
 XX
 PN DE19915485-A1.
 XX
 PD 19-OCT-2000.
 XX
 PF 07-APR-1999; 99DE-01015485.
 XX
 PR 07-APR-1999; 99DE-01015485.
 XX
 PA (KATU/) KATUS H A.
 PA (REMP/) REMPPIS A.
 XX
 PI Katus HA, Remppis A;
 XX
 DR WPI; 2000-673510/66.
 DR N-PSDB; AAC81804.
 XX
 PT Composition containing S100 protein, corresponding nucleic acid or
 PT vector, useful for treating cardiomyopathy and cardiac insufficiency.
 XX
 PS Claim 35; Page 11; 36pp; German.
 XX
 CC This invention describes a novel composition for treating primary or
 CC secondary cardiomyopathy or cardiac insufficiency contains at least one
 CC S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or
 CC fragments, or a gene transfer vector containing (II), optionally
 CC formulated with auxiliaries and/or carriers. (I) are calcium-binding
 CC proteins involved in calcium homeostasis, so their overexpression in
 CC cardiac muscle will improve pumping capacity (and overall capacity) of
 CC the heart. In cultured myocardial cells they increase the contraction and
 CC relaxation rates associated with increased systolic calcium ion release
 CC from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are
 CC used to treat cardiomyopathy (CMP) where inherited or caused by

CC spontaneous mutations and ischemic CMP caused by arteriosclerosis,
 CC dilative CMP caused by toxic/infectious disease, cardiac disease caused
 CC by pulmonary and/or arterial hypertension, and structural disease caused
 CC by rhythm disorders or valve defects, generally any condition associated
 CC with reduced contractile force. Unlike calmodulin, which is expressed
 CC ubiquitously, (1) show tissue-specific expression and treat the
 CC underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac
 CC disease
 CC
 CC Sequence 101 AA;
 SQ
 Query Match 100.0%; Score 65; DB 3; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.00023;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LPSFLGKRTDEAA 13
 |||||
 Db 42 LPSFLGKRTDEAA 54
 |||||
 RESULT 7
 AAB37432
 ID AAB37432 standard; protein; 101 AA.
 XX
 AC AAB37432;
 XX
 DT 21-FEB-2001 (first entry)
 XX
 DE Human mts-1.
 XX
 KW Human; mts-1; cytostatic; cancer; metastasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200064475-A1.
 XX
 PD 02-NOV-2000.
 XX
 PF 20-APR-2000; 2000WO-US011006.
 XX
 PR 23-APR-1999; 99US-00298625.
 XX
 PA (RESE) RESEARCH CORP TECHNOLOGIES INC.
 XX
 PI Lukanidin E;
 XX
 DR WPI; 2000-687266/67.
 XX
 DR N-PSDB; AAC68131, AAC68132.
 XX
 XX Treating cancer and preventing metastasis comprises administration of an
 PT anti-mts-1 protein antibody or antisense oligonucleotide.
 XX
 PS Example 6; Fig 2; 155pp; English.
 XX
 CC The present sequence is human mts-1. Mts-1 protein is a calcium-binding
 CC protein, and is thought to have a role in myoepithelial cell
 CC differentiation. The present invention relates to methods for treating
 CC cancer and preventing metastases, comprising the administration of a
 CC composition directed against the mts-1 protein
 CC
 CC Sequence 101 AA;
 SQ
 Query Match 100.0%; Score 65; DB 3; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.00023;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LPSFLGKRTDEAA 13
 |||||
 Db 42 LPSFLGKRTDEAA 54
 |||||
 RESULT 8
 AAB72387

ID AAB72387 standard; protein; 101 AA.
 XX
 AC AAB72387;
 XX
 DT 24-MAY-2001 (first entry)
 XX
 DE Murine Mts 1 protein amino acid sequence.
 XX
 KW Mts 1; Mts1-del75; Mts1-4S; S100A4; neuronal cell growth; cardiac arrest;
 KW neuronal degeneration; Parkinson's disease; Alzheimer's disease;
 KW Down's syndrome; stroke; sciatic crush; spinal cord injury; mouse.
 XX
 OS Mus sp.
 XX
 PN WO200118043-A2.
 XX
 PD 15-MAR-2001.
 XX
 PF 07-SEP-2000; 2000WO-US024495.
 XX
 PR 10-SEP-1999; 99US-00393433.
 XX
 PA (PROL-) PROLIFIA INC.
 XX
 PI Bock E, Lukanidin EM, Berezin V;
 XX
 DR WPI; 2001-235188/24.
 XX
 XX New isolated functional derivatives of Mts protein for stimulating
 PT neuronal cell growth to treat neurological conditions e.g. Parkinson's
 PT disease, Alzheimer's disease, Down's syndrome and stroke.
 XX
 PS Disclosure; Page; 60pp; English.
 XX
 CC This invention relates to functional derivatives of an Mts protein,
 CC particularly Mts1-del75, and Mts1-4S. Mts 1 (also known as S100A4) is a
 CC member of the S100 gene family. Mts1 proteins and their functional
 CC derivatives are used to stimulate the growth of neuronal cells and in the
 CC treatment of neurological conditions including those characterised by
 CC neuronal degeneration, death or injury such as Parkinson's disease,
 CC Alzheimer's disease, Down's syndrome, stroke, cardiac arrest, sciatic
 CC crush, spinal cord injury, injury to sensory neurons or degenerative
 CC disease of the retina. The present sequence represents the murine Mts 1
 CC protein, from which the Mts proteins of the invention may be derived.
 CC NOTE: The present sequence is not shown in the specification but it has
 CC been derived from the sequence given in Genes Dev. 3, 1989 (referred to
 CC on page 10 of the specification)
 XX
 XX Sequence 101 AA;
 SQ
 Query Match 100.0%; Score 65; DB 4; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.00023;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LPSFLGKRTDEAA 13
 |||||
 Db 42 LPSFLGKRTDEAA 54
 |||||
 RESULT 9
 AAB72386
 ID AAB72386 standard; protein; 101 AA.
 XX
 AC AAB72386;
 XX
 DT 24-MAY-2001 (first entry)
 XX
 DE Human Mts 1 protein amino acid sequence.
 XX
 KW Mts 1; Mts1-del75; Mts1-4S; S100A4; neuronal cell growth; cardiac arrest;
 KW neuronal degeneration; Parkinson's disease; Alzheimer's disease;
 KW Down's syndrome; stroke; sciatic crush; spinal cord injury; human.
 XX

OS Homo sapiens.
 XX WO200118043-A2.
 XX
 XX PD 15-MAR-2001.
 XX PF 07-SEP-2000; 2000WO-US024495.
 XX PR 10-SEP-1999; 99US-00393433.
 XX PA (PROL-) PROLIFIA INC.
 XX PI Bock E, Lukanidin EM, Berezin V;
 XX WPI; 2001-235186/24.
 XX DR
 XX PT New isolated functional derivatives of Mts protein for stimulating
 PT neuronal cell growth to treat neurological conditions e.g. Parkinson's
 PT disease, Alzheimer's disease, Down's syndrome and stroke.
 XX PS Disclosure; Page; 60pp; English.
 XX CC This invention relates to functional derivatives of an Mts protein,
 CC particularly Mts1-del175, and Mts1-48. Mts1 (also known as S100A4) is a
 CC member of the S100 gene family. Mts1 proteins and their functional
 CC derivatives are used to stimulate the growth of neuronal cells and in the
 CC treatment of neurological conditions including those characterised by
 CC neuronal degeneration, death or injury such as Parkinson's disease,
 CC Alzheimer's disease, Down's syndrome, stroke, cardiac arrest, sciatic
 CC crush, spinal cord injury, injury to sensory neurons or degenerative
 CC disease of the retina. The present sequence represents the human Mts 1
 CC protein, from which the Mts proteins of the invention may be derived.
 CC NOTE: The present sequence is not shown in the specification but it has
 CC been derived from the sequence given in US patent 5801142 (referred to on
 XX page 10 of the specification)
 XX SQ Sequence 101 AA;
 Query Match 100.0%; Score 65; DB 4; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.00023;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LPSFLGKRTDEAA 13
 |||||
 Db 42 LPSFLGKRTDEAA 54
 RESULT 10
 ABU08513
 ID ABU08513 standard; protein; 101 AA.
 XX AC ABU08513;
 XX DT 27-MAY-2003 (first entry)
 XX DE Human mts-1 protein.
 XX KW Human; mts-1; metastatic cancer; p53; cytostatic; gene therapy.
 XX OS Homo sapiens.
 XX PN US2002172680-A1.
 XX PD 21-NOV-2002.
 XX PF 29-APR-2002; 2002US-00135152.
 XX PR 09-JUL-1990; 90US-00550600.
 XX PR 25-NOV-1992; 92US-00981455.
 XX PR 31-JAN-1994; 94US-00190560.
 XX PR 06-JUN-1995; 95US-00468942.
 XX PR 23-APR-1999; 99US-00298635.
 XX

PA (LUKA/) LUKANIDIN E.
 XX Lukanidin E;
 XX PI
 XX WPI; 2003-328422/31.
 XX DR N-PSDB; ABX93578.
 XX PT
 PT Treating cancer by administering a reagent directed against the mts-1
 PT protein or an oligonucleotide capable of binding to mts-1 mRNA.
 XX PS Example 6; Fig 2; 64pp; English.
 XX CC The invention relates to treating cancer comprising administering a
 CC reagent directed against the mts-1 (not defined) protein or an
 CC oligonucleotide capable of binding to mts-1 mRNA. Also include are a
 CC method of inactivating, destroying or nullifying a mts-1 protein or cells
 CC and a method of inhibiting metastasis in a cancerous cell. Experiments
 CC showed that mts-1 inhibited the phosphorylation of full-size p53 and the
 CC C-terminal protein fragment by PKC. Addition of the same concentrations
 CC of mts-1 to the PKC reaction mixture did not affect the phosphorylation
 CC of the N-terminal and DNA-binding domains of p53. The method is useful
 CC for treating cancers associated with the mts-1 gene. The present sequence
 CC represents human mts-1 protein
 XX SQ Sequence 101 AA;
 Query Match 100.0%; Score 65; DB 6; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.00023;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 LPSFLGKRTDEAA 13
 |||||
 Db 42 LPSFLGKRTDEAA 54
 RESULT 11
 ADD14157
 ID ADD14157 standard; protein; 101 AA.
 XX AC ADD14157;
 XX DT 01-JAN-2004 (first entry)
 XX DE Human src biomarker polypeptide SEQ ID NO:346.
 XX KW predictor set; protein tyrosine kinase activity modulator;
 KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
 KW gene therapy; drug sensitivity; genetic profile; cancer; human.
 XX OS Homo sapiens.
 XX PN WO2003062395-A2.
 XX PD 31-JUL-2003.
 XX PF 17-JAN-2003; 2003WO-US001981.
 XX PR 18-JAN-2002; 2002US-0350061P.
 XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX PI Huang F, Fairchild CR, Lee FY, Shaw P;
 XX WPI; 2003-636735/60.
 XX N-PSDB; ADD14760.
 XX PT New polynucleotides and polypeptides for predicting the activity of
 PT compounds that interact with protein tyrosine kinases and/or protein
 XX tyrosine kinase pathways.
 XX PS Claim 10; SEQ ID NO 346; 139pp; English.
 XX CC The present invention describes a predictor set comprising a plurality of

CC polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.

XX
SQ Sequence 101 AA;

Query Match 100.0%; Score 65; DB 7; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.00023; Mismatches 0; Gaps 0;
Matches 13; Conservative 0;

Qy 1 LPSFLGKRTDEAA 13
Db 42 LPSFLGKRTDEAA 54
|||||

RESULT 12
ABB93293
ID ABB93293 standard; protein; 521 AA.
XX
AC ABB93293;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 2504.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX
PA (FARB) BAYER AG.
XX
XX Tietjen K, Weidler M;
XX WPI; 2002-269010/31.
XX
XX Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX
PS Claim 5; SEQ ID NO 2504; 261pp + Sequence Listing; English.
XX

CC The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides

XX
SQ Sequence 521 AA;

Query Match 67.7%; Score 44; DB 5; Length 521;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 LPSFLGKRTDEA 12
Db 67 LPSYIGKRMDES 78
|||||

RESULT 13
AAR27058
ID AAR27058 standard; protein; 97 AA.
XX
AC AAR27058;

DT 25-MAR-2003 (revised)
DT 25-FEB-1993 (first entry)
XX
DE Sequence of small Ca++ binding proteins encoded by Can19 clone 19.
XX
KW Can19; tumour suppressor gene; cancer; therapy.

XX Homo sapiens.
XX WO9215602-A1.
XX 17-SEP-1992.
XX 28-FEB-1992; 92WO-US001624.
XX 28-FEB-1991; 91US-00662216.
XX (DAND) DANA FARBER CANCER INST INC.
XX Sager R;

WPI; 1992-331663/40.
N-PSDB; AAQ28760.
Diagnosis and treatment of cancer - using candidate tumour suppressor
genes or the corresp. antibodies.
Disclosure; Page 31; 54pp; English.

A clone originally termed clone 19, and now referred to as Can19,
represents a gene expressed in normal mammary epithelial cell strains but
not in tumor-derived cell lines. Sequence comparisons have shown that
Can19 is a member of the S100 gene family, encoding small Ca++ binding
proteins (about 10 kD) with diverse functions. These proteins have two
"EF hands", domains where Ca2+ is bound. Can19 is also related in
structure to the small regulatory subunit of calpactin, p11. Can19 is not
expressed in breast tumor cells. Can19 appears to be negatively regulated
in tumors, in contrast to calyculin. (Updated on 25-MAR-2003 to correct
PN field.) (Updated on 25-MAR-2003 to correct PA field.)

XX
SQ Sequence 97 AA;

Query Match 64.6%; Score 42; DB 2; Length 97;
Best Local Similarity 63.6%; Pred. No. 5.3;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDE 11
 Db 42 LPSFVGKVD 52

RESULT 14

AA26406
 ID AAR26406 standard; protein; 97 AA.
 XX AC AAR26406;
 XX DT 25-MAR-2003 (revised)
 DT 27-FEB-1993 (first entry)
 XX DE Sequence of the clone 19 gene product.
 XX KW Clone 19; diagnosis; prognosis; cancer; tumour.
 XX OS Homo sapiens.
 XX PN WO9215600-A1.
 XX PD 17-SEP-1992.
 XX PF 28-FEB-1992; 92WO-US001625.
 XX PR 28-FEB-1991; 91US-00662198.
 XX PA (DAND) DANA FARBER CANCER INST INC.
 XX PI Sager R, Lee SW, Tomasetto C;
 XX WPI; 1992-331662/40.
 XX DR N-PSDB; AAQ28007.
 XX PT Clone 19 gene prodn. and its DNA and antibody - for diagnosis, prognosis
 PT and treatment of solid tumours, especially breast cancer.
 XX PS Disclosure; Page 12-13; 20pp; English.
 XX CC Clone 19 was derived from normal human mammary epithelial cells strain
 CC 76N. Clone 19 represents a gene expressed in normal mammary epithelial
 CC cell strains but not in tumour-derived cell lines. Sequence comparisons
 CC have shown that it is a member of the S100 gene family, encoding small
 CC Ca++ binding proteins (about 10kd) with diverse functions. These proteins
 CC have two 'EF hands', domains where Ca++ is bound, in contrast to calmodulin
 CC proteins which have four. Clone 19 is also related in structure to the
 CC small regulatory subunit of calpactin, p11. MRP8 and MRP14 are also
 CC related, and are S100 proteins expressed by macrophages during chronic
 CC inflammation. (Updated on 25-MAR-2003 to correct PN field.) (Updated on
 CC 25-MAR-2003 to correct PA field.)
 XX SQ Sequence 97 AA;

Query Match 64.6%; Score 42; DB 2; Length 97;
 Best Local Similarity 63.6%; Pred. No. 5.3;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDE 11
 Db 42 LPSFVGKVD 52

RESULT 15

AA45532
 ID AAB45532 standard; protein; 97 AA.
 XX AC AAB45532;
 XX DT 22-FEB-2001 (first entry)
 XX DE Human S100A2 protein.

KW S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;
 KW calcium-binding protein; calcium homeostasis; cardiac muscle;
 KW pumping capacity; myocardial cell; systolic calcium ion release;
 KW sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;
 KW valve defect.

XX OS Homo sapiens.
 XX PN DE19915485-A1.
 XX PD 19-OCT-2000.
 XX PF 07-APR-1999; 99DE-01015485.
 XX PR 07-APR-1999; 99DE-01015485.
 XX PA (KATU/) KATUS H A.
 XX PA (REMP/) REMPPIS A.
 XX PI Katus HA, Remppis A;
 XX WPI; 2000-673510/66.
 XX CC Composition containing S100 protein, corresponding nucleic acid or
 CC vector, useful for treating cardiomyopathy and cardiac insufficiency.
 XX PS Claim 35; Page 9; 36pp; German.

XX CC This invention describes a novel composition for treating primary or
 CC secondary cardiomyopathy or cardiac insufficiency contains at least one
 CC S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or
 CC fragments, or a gene transfer vector containing (II), optionally
 CC formulated with auxiliaries and/or carriers. (I) are calcium-binding
 CC proteins involved in calcium homeostasis, so their overexpression in
 CC cardiac muscle will improve pumping capacity (and overall capacity) of
 CC the heart. In cultured myocardial cells they increase the contraction and
 CC relaxation rates associated with increased systolic calcium ion release
 CC from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are
 CC used to treat cardiomyopathy (CMP) where inherited or caused by
 CC spontaneous mutations and ischemic CMP caused by arteriosclerosis,
 CC dilative CMP caused by toxic/infectious disease, cardiac disease caused
 CC by pulmonary and/or arterial hypertension, and structural disease caused
 CC by rhythm disorders or valve defects, generally any condition associated
 CC with reduced contractile force. Unlike calmodulin, which is expressed
 CC ubiquitously, (I) show tissue-specific expression and treat the
 CC underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac
 CC disease

XX SQ Sequence 97 AA;

Query Match 64.6%; Score 42; DB 3; Length 97;
 Best Local Similarity 63.6%; Pred. No. 5.3;
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 LPSFLGKRTDE 11
 Db 42 LPSFVGKVD 52

Search completed: September 9, 2004, 10:53:06
 Job time : 29.3419 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:42:57 ; Search time 26.6323 Seconds
(without alignments)
189.556 Million cell updates/sec

Title: US-10-067-618-5

Perfect score: 80

Sequence: 1 KEGDKFKLNKSELKEL 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	77	96.2	101	6	Q9TV56
2	65	81.2	94	11	Q9D3M4
3	65	81.2	94	11	Q91V77
4	64	80.0	92	11	Q925T3
5	61	76.2	79	11	Q9JL08
6	58	72.5	89	11	Q9R2B7
7	54	67.5	756	16	Q8REN0
8	49	61.3	92	13	Q9PSF6
9	49	61.3	405	16	Q8RDT1
10	48	60.0	338	5	Q8ISK7
11	47	58.8	99	13	Q8AYJ2
12	46	57.5	277	16	Q9X204
13	46	57.5	767	13	Q9OWS0
14	45	56.2	101	13	Q93395
15	45	56.2	201	16	Q892B0
16	45	56.2	353	16	Q892H4

17	44.5	55.6	58	17	029879
18	44.5	55.6	1086	16	Q7VK42
19	44	55.0	79	6	Q865V3
20	44	55.0	80	11	Q91XG5
21	44	55.0	216	16	Q93M84
22	44	55.0	478	13	Q93313
23	44	55.0	478	13	Q91AM4
24	43	53.8	227	16	Q98RK4
25	43	53.8	311	17	Q8TW68
26	43	53.8	398	4	Q8NA87
27	43	53.8	480	9	Q9AZE5
28	43	53.8	480	16	Q9CI49
29	43	53.8	510	4	Q8WZ0
30	43	53.8	511	4	Q8WZ0
31	43	53.8	1307	16	Q8FIP0
32	43	53.8	1307	16	Q8K792
33	43	53.8	3377	5	Q8IEH5
34	42.5	53.1	339	16	Q55871
35	42	52.5	176	17	Q58049
36	42	52.5	302	10	Q8GZ11
37	42	52.5	341	17	Q8TSC0
38	42	52.5	463	12	Q85291
39	42	52.5	463	12	Q91V54
40	42	52.5	475	2	Q59956
41	42	52.5	478	16	Q7V2K3
42	42	52.5	488	3	Q9C222
43	42	52.5	500	12	Q91LB9
44	42	52.5	501	12	Q8VAX1
45	42	52.5	728	16	Q8XK95

ALIGNMENTS

RESULT 1

Q9TV56
ID Q9TV56 PRELIMINARY; PRT; 101 AA.
AC Q9TV56
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Metastasin.
GN MTS1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madin-Darby; TISSUE=Kidney;
RA Miyamori H., Hasegawa K., Kim K., Sato H.;
RT "Expression of metastasin associated mts1 gene is co-induced with
RT membrane type-1 matrix metalloproteinase (MT1-MMP) during oncogenic
RT transformation and tubular formation of madin darby canine kidney
RT (MDCK) epithelial cells";
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AB031064; BAA83419.1; -.
DR HSSP; P30801; 1A03.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD0003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CBP; 1.
SQ SEQUENCE 101 AA; 11847 MW; 5535387DB7577DF0 CRC64;

Query Match 96.2%; Score 77; DB 6; Length 101;
Best Local Similarity 93.8%; Pred. No. 0.00028;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;


```

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AK012578; BAB28330.1; -.
DR EMBL; BC005590; AAH05590.1; -.
DR EMBL; AF368423; AAL14436.1; -.
DR EMBL; AK002721; BAB22308.1; -.
DR MGD; MGI:1338917; S100a1.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00015; EF-hand; 1.
DR PROSITE; PS00018; EF-hand; 1.
DR PROSITE; PS00303; S100_CaBP; 1.
SQ SEQUENCE 94 AA; 10505 NW; 7CBE817BBF36ED15 CRC64;

Query Match 81.2%; Score 65; DB 11; Length 94;
Best Local Similarity 75.0%; Pred. No. 0.018;
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KEGDKFKLNKSELKEL 16
Db 22 QEGDKYKLSKKELKDL 37

RESULT 4
Q925T3 ID Q925T3 PRELIMINARY; PRT; 92 AA.
AC Q925T3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S100B.
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RA Kuge O., Yamakawa Y., Nishijima M.;
RT "Enhancement of transport-dependent decarboxylation of
RT phosphatidylserine by S100B protein in permeabilized Chinese hamster
RT ovary cells.";
RL J. Biol. Chem. 0:0-0(2001).
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AB056121; BAB43945.1; -.
DR GO; GO:0005737; Cytoplasm; ISS.
DR GO; GO:0005576; Extracellular; ISS.
DR GO; GO:0005509; F:calcium ion binding; ISS.
DR GO; GO:0019210; F:kinase inhibitor activity; ISS.
DR GO; GO:0042803; F:protein homodimerization activity; ISS.
DR GO; GO:0048155; F:S100 alpha binding; ISS.
DR GO; GO:0048154; F:S100 beta binding; ISS.
DR GO; GO:0048156; F:tau protein binding; ISS.
DR GO; GO:0008270; F:zinc ion binding; ISS.
DR GO; GO:0048143; F:astrocyte activation; ISS.
DR GO; GO:0007409; F:axonogenesis; ISS.
DR GO; GO:0006874; F:calcium ion homeostasis; ISS.
DR GO; GO:0006112; F:energy reserve metabolism; ISS.
DR GO; GO:0048151; F:hyperphosphorylation; ISS.
DR GO; GO:0006917; P:induction of apoptosis; ISS.
DR GO; GO:0007611; P:learning and/or memory; ISS.
DR GO; GO:0045917; P:positive regulation of complement activation; ISS.
DR GO; GO:0042035; P:regulation of cytokine biosynthesis; ISS.
DR GO; GO:0048169; P:regulation of long-term neuronal synaptic p. . .; ISS.
DR GO; GO:0006417; P:regulation of protein biosynthesis; ISS.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.

```

```

DR ProDom; PD003407; CaBP_S100; 1.
DR PROSITE; PS00018; EF-hand; 1.
DR PROSITE; PS00303; S100_CaBP; 1.
SQ SEQUENCE 92 AA; 10749 NW; AF50107EC2BEDF6B CRC64;

Query Match 80.0%; Score 64; DB 11; Length 92;
Best Local Similarity 81.2%; Pred. No. 0.025;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KEGDKFKLNKSELKEL 16
Db 21 QEGDKYKLSKKELKDL 36

RESULT 5
Q9JL08 ID Q9JL08 PRELIMINARY; PRT; 79 AA.
AC Q9JL08;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S100 calcium binding protein A1 (Fragment).
GN S100A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kiewitz R., Lyons G.E., Schafer B.W., Heizmann C.W.;
RT "Transcriptional regulation of S100A1 and expression during mouse
RT heart development.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AF218353; AAF32320.1; -.
DR HSP; P04631; IB4C.
DR MGD; MGI:1338917; S100a1.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF-hand; 1.
DR PROSITE; PS00303; S100_CaBP; 1.
DR NON TER 79
FT NON TER 79
SQ SEQUENCE 79 AA; 8863 MW; F94EDA3A798815D5 CRC64;

Query Match 76.2%; Score 61; DB 11; Length 79;
Best Local Similarity 68.8%; Pred. No. 0.063;
Matches 11; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KEGDKFKLNKSELKEL 16
Db 22 QEGDKYKLSKKELKDL 37

RESULT 6
Q9R2B7 ID Q9R2B7 PRELIMINARY; PRT; 89 AA.
AC Q9R2B7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Calyculin (Calcium binding protein).
GN CACY OR S100A6.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=WISTAR;
RA Konrad L., Gabius H.J., Amueller G.;
RT "Sequence and expression study of calcyclin in the rat testis.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RA Ito M., Kizawa K.;
RT "Expression of S100 genes in hair follicle epithelium.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AJ132717; CAB24202.1; -;
DR EMBL; AF140232; AAK28306.1; -;
DR HSPG; P30801; 1A03.
DR GO; GO:0005635; C:nuclear membrane; ISS.
DR GO; GO:0005509; F:calcium ion binding; ISS.
DR GO; GO:0005515; F:protein binding; ISS.
DR GO; GO:0007409; P:angiogenesis; ISS.
DR GO; GO:0004814; P:positive regulation of fibroblast proliferation. . .; ISS.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CABP_S100; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Cyclin.
SQ SEQUENCE 89 AA; 10035 MW; 2AALA4163D57DC87 CRC64;

Query Match 72.5%; Score 58; DB 11; Length 89;
Best Local Similarity 75.0%; Pred. No. 0.2; Mismatches 3; Indels 0; Gaps 0;
Matches 12; Conservative 1;

QY 1 KEGDKFKLNKSELKEL 16
DB 22 KEGDKHTLSKKELKEL 37
||||| :|:|||||

RESULT 7
Q8RENO PRELIMINARY; PRT; 756 AA.
AC Q8RENO;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE DNA topoisomerase I (EC 5.99.1.2).
GN FN1069.
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=21886394; PubMed=11889109;
RA Kapral V., Anderson I., Ivanova N., Renzik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Fongstein M., Kyripides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018 (2002).
DR EMBL; AE010612; AAL95265.1; -.
DR GO; GO:0005694; C:chromosome; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003917; F:DNA topoisomerase type I activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006304; P:DNA modification; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR GO; GO:0006268; P:DNA unwinding; IEA.
DR InterPro; IPR003601; DNATopi ATP bind.
DR InterPro; IPR003602; DNATopi_DNA_bind.

```

```

DR InterPro; IPR005733; DNA_topI_bact.
DR InterPro; IPR000380; DNA_topoisomerase.
DR InterPro; IPR006171; Toprim_dom.
DR InterPro; IPR006154; Toprim_sub.
DR Pfam; PF01131; Topoisom_bac; 1.
DR Pfam; PF01751; Toprim; 1.
DR Pfam; PF01396; zf-C4_Topoiso; 1.
DR PRINTS; PR00417; PRTFISMRASEI.
DR SMART; SM00437; TOP1AC; 1.
DR SMART; SM00436; TOP1BC; 1.
DR SMART; SM00493; TOPRIM; 1.
DR TIGREMS; TIGR01051; topa_bact; 1.
DR PROSITE; PS00396; TOPOISOMERASE_I_PROK; 1.
KW Isomerase; Complete proteome.
SQ SEQUENCE 756 AA; 87244 MW; 384DB1617F7FB106 CRC64;

Query Match 67.5%; Score 54; DB 16; Length 756;
Best Local Similarity 66.7%; Pred. No. 6.5;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKE 15
DB 438 KEGDKFTLDKLDIKE 452
||||| :|:|||||

RESULT 8
Q9PSF6 PRELIMINARY; PRT; 92 AA.
AC Q9PSF6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Ictacalcin.
OS Ictalurus punctatus (Channel catfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;
OC Ictaluridae; Ictalurus.
OX NCBI_TaxID=7998;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94376615; PubMed=8090068;
RA Bettini E., Porta A.R., Dahmen N., Wang H., Margolis F.L.;
RT "Expressed sequence tags (EST) identify genes preferentially expressed
in catfish chemosensory tissues.";
RL Brain Res. Mol. Brain Res. 23:285-291 (1994).
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR HSP; P30801; 1A03.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CABP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
DR PROSITE; PS00303; S100_CABP; 1.
SQ SEQUENCE 92 AA; 10022 MW; E04875D0C9921C50 CRC64;

Query Match 61.3%; Score 49; DB 13; Length 92;
Best Local Similarity 62.5%; Pred. No. 5;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
DB 21 EGDKCTLTGKELKDL 36
||||| :|:|||||

RESULT 9
Q8RDT1 PRELIMINARY; PRT; 405 AA.
ID Q8RDT1
AC Q8RDT1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Flavoprotein.
GN FN1423
OS Fusobacterium nucleatum (subsp. nucleatum).
OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;
OC Fusobacterium.
OX NCBI_TaxID=76856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25586;
RX MEDLINE=2186394; PubMed=11889109;
RA Kapatal V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,
RA Bhattacharya A., Bartman A., Gardner W., Grechkin G., Zhu L.,
RA Vasileva O., Chu L., Kogan A., Chaga O., Goltzman E., Bernal A.,
RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,
RA Forstein M., Kyrpides N., Overbeek R.;
RT "Genome sequence and analysis of the oral bacterium Fusobacterium
RT nucleatum strain ATCC 25586."
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE010647; AAL95616.1; -.
DR InterPro; IPR001279; Bactnase-like.
DR InterPro; IPR008254; Flav_nitox_synth.
DR Pfam; PF00258; Flavodoxin; 1.
DR Pfam; PF00753; lactamase B; 1.
DR PROSITE; PS05092; FLAVODOXIN_LIKE; 1.
KW Complete proteome.
SQ SEQUENCE 405 AA; 46641 MW; A7F666C1A56A881FA CRC64;

Query Match 61.3%; Score 49; DB 16; Length 405;
Best Local Similarity 69.2%; Pred. No. 21;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSEL 13
DB 128 KEGDKFKFGKHEI 140
|:|||||:|:|

RESULT 10
Q815K7 PRELIMINARY; PRT; 338 AA.
AC Q815K7
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN Pf0990W.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Anguoli S.,
RA Partea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairhead A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AE014847; AAN36284.1; -.
KW Hypothetical protein.
SQ SEQUENCE 338 AA; 40678 MW; 39930E75788725B7 CRC64;

Query Match 60.0%; Score 48; DB 5; Length 338;
Best Local Similarity 62.5%; Pred. No. 25;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```

```

QY 1 KEGDKFKLNKSELKEL 16
|:|||||:|:|
DB 69 KDTDRFKLNKEKEYL 84

RESULT 11
Q8AYJ2 PRELIMINARY; PRT; 99 AA.
AC Q8AYJ2
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S-100 calcium-binding protein A1.
OS Squalus acanthias (Spiny dogfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squala; Squaloidei; Squalidae; Squalus.
OX NCBI_TaxID=7797;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang C., Callard G.V.;
RT "Molecular cloning and stage dependence of an S-10 cDNA from the shark
RT testis."
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF421551; AAN63527.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR SMART; SM00054; EFh; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CaBP; 1.
SQ SEQUENCE 99 AA; 11050 MW; BA62D8190A4A3693 CRC64;

Query Match 58.8%; Score 47; DB 13; Length 99;
Best Local Similarity 50.0%; Pred. No. 11;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
|:|||||:|:|
DB 21 KEGDKYSLNNEMVDL 36

RESULT 12
Q9X204 PRELIMINARY; PRT; 277 AA.
AC Q9X204
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein TM1676.
GN TM1676.
OS Thermotoga maritima.
OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."
RL Nature 399:323-329(1999).
DR EMBL; AE001808; AAD36743.1; -.
DR PIR; B72226; B72226.
DR TIGR; TM1676; -.

```

```
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 277 AA; 32569 MW; F8F5D215954A835D CRC64;

Query Match 57.5%; Score 46; DB 16; Length 277;
Best Local Similarity 71.4%; Pred. No. 42;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 GDKFKLNKSELKEL 16
Db 238 GKGLGINKKELKEL 251

RESULT 13
Q90WS0 PRELIMINARY; PRT; 767 AA.
AC Q90WS0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Topoisomerase I alpha.
GN TOPIALPHA.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Smith S.F., Metcalfe J., Elgar G.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035357; CAA23015.1; -.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003917; F:DNA topoisomerase type I activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR GO; GO:0006268; P:DNA unwinding; IEA.
DR InterPro; IPR001631; Topoisomerase_I.
DR InterPro; IPR008336; Topoisomerase_I_N.
DR Pfam; PF01028; Topoisomerase_I; 1.
DR Pfam; PF02919; Topoisomerase_I_N; 1.
DR PRINTS; PR00416; EUTPISMRASEI.
DR SMART; SM00435; TOPEUG; 1.
DR PROSITE; PS00176; TOPOISOMERASE_I_EUK; 1.
KW Isomerase.
SQ SEQUENCE 767 AA; 90821 MW; 50018E2E1EF97D00 CRC64;

Query Match 57.5%; Score 46; DB 13; Length 767;
Best Local Similarity 66.7%; Pred. No. 11e+02;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKE 15
Db 42 KERDKFKHNSEHKD 56

RESULT 14
O93395 PRELIMINARY; PRT; 101 AA.
AC O93395;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S100-like calcium binding protein.
GN S100.
OS Salvelinus fontinalis (Brook trout) (Brook char).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8038;
RN [1]
RP SEQUENCE FROM N.A.
TTISSUE=Ovulatory ovary;
```


GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:32:57 ; Search time 4.54194 Seconds
(without alignments)
183.429 Million cell updates/sec

Title: US-10-067-618-5
Perfect score: 80
Sequence: 1 KEGDKFKLNSKELXEL 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	100	1 S104_BOVIN	P35466 bos taurus
2	80	100.0	101	1 S104_HUMAN	P26447 homo sapien
3	77	96.2	101	1 S104_MOUSE	P07091 mus musculus
4	72	90.0	101	1 S104_RAT	P05942 rattus norv
5	68	85.0	93	1 S10A_BOVIN	P02639 bos taurus
6	68	85.0	93	1 S10A_HUMAN	P23297 homo sapien
7	65	81.2	93	1 S10A_RAT	P35467 rattus norv
8	65	81.2	97	1 S102_BOVIN	P10462 bos taurus
9	65	81.2	97	1 S102_HUMAN	P29034 homo sapien
10	64	80.0	91	1 S10B_BOVIN	P02638 bos taurus
11	64	80.0	91	1 S10B_HUMAN	P04271 homo sapien
12	64	80.0	91	1 S10B_MOUSE	P50114 mus musculus
13	64	80.0	91	1 S10B_RAT	P04631 rattus norv
14	61	76.2	93	1 S10A_MOUSE	P56565 mus musculus
15	58	72.5	89	1 S106_MOUSE	P14069 mus musculus
16	58	72.5	90	1 S106_RABIT	P30801 oryctolagus
17	58	72.5	90	1 S106_RAT	P05964 rattus norv
18	57	71.2	92	1 S106_CHICK	P07953 gallus gall
19	55	68.8	90	1 S106_HUMAN	P06703 homo sapien
20	53	66.2	92	1 S106_HORSE	O77691 equus caball
21	53	66.2	92	1 S10I_ICTPU	Q91061 ictalurus p
22	52	65.0	101	1 S103_MOUSE	P56566 mus musculus
23	51	63.7	101	1 S103_HUMAN	P33764 homo sapien
24	49	61.3	78	1 S10D_PIG	P02632 sus scrofa
25	47	58.8	92	1 S105_HUMAN	O88945 mus musculus
26	47	58.8	93	1 S105_MOUSE	O67066 aquifex aeo
27	46.5	58.1	461	1 FTSY_AQUAE	P32908 saccharomyc
28	45	56.2	1225	1 SMCL1_YEAST	P02633 bos taurus
29	44	55.0	78	1 S10D_BOVIN	P29377 homo sapien
30	44	55.0	78	1 S10D_HUMAN	P97816 mus musculus
31	44	55.0	78	1 S10D_MOUSE	P02634 rattus norv
32	44	55.0	78	1 S10D_RAT	O27123 methanobact
33	44	55.0	519	1 RPB2_METTH	

34 43.5 54.4 328 1 YD70 METUA
35 43 53.8 1173 1 ATC2 YEAST
36 43 53.8 1333 1 PAD3 MOUSE
37 42 52.5 478 1 SYE PROMP
38 42 52.5 490 1 C72T ARATH
39 42 52.5 621 1 HTPG_HELPJ
40 42 52.5 623 1 HTPG_CLOPE
41 42 52.5 1030 1 ATHA XENLA
42 41 51.2 75 1 R17E THEAC
43 41 51.2 101 1 S111 CHICK
44 41 51.2 119 1 M126 CHICK
45 41 51.2 147 1 YXND_EACSU

Q58765 methanococ
P38929 saccharomyc
Q99ht2 mus musculu
Q7v2k3 prochloroco
Q9sae4 arabidopsis
Q9znm2 helicobacte
Q8xnc2 clostridium
Q92126 xenopus lae
Q9hkk8 thermoplas
P24479 gallus gall
P28318 gallus gall
P40737 bacillus su

ALIGNMENTS

RESULT 1
S104_BOVIN
ID S104_BOVIN STANDARD; PRT; 100 AA.
AC P35466;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Placental calcium-binding protein homolog.
GN S100A4 OR CAPL.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Retina;
RX MEDLINE=94164991; PubMed=8119967;
RA Polans A.S., Palczewski K., Asson-Batres M.A., Ohguro H., Witowska D.,
RA Haley T.L., Baizer L., Crabb J.W.;
RT "Purification and primary structure of Capl, an S-100-related
RT calcium-binding protein isolated from bovine retina.";
RL J. Biol. Chem. 269:6233-6240(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98430466; PubMed=9759666;
RA Duarte W.R., Kasugai S., Iimura T., Oida S., Takenaga K., Ohya K.,
RA Ishikawa I.;
RT "cDNA cloning of S100 calcium-binding proteins from bovine
RT periodontal ligament and their expression in oral tissues.";
RL J. Dent. Res. 77:1694-1699(1998).
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D89056; BAA13754.1; --
DR PIR: A53217; A53217.
DR HSSP: P30801; IA03.
DR InterPro: IPR001751; CaBP_S100.
DR InterPro: IPR002048; EF-HAND.
DR Pfam: PF00036; ehand; 1.
DR Pfam: PF01023; S100; 1.
DR ProDom: PD003407; CaBP_S100; 1.
DR ProDom: PD000012; EF-hand; 1.
DR PROSITE: PS00018; EF-HAND; 1.
DR PROSITE: PS00303; S100_CaBP; 1.
DR Calcium-binding; Acetylation.
FT INIT MET 0
FT MOD_RES 1 1 ACETYLATION.

```

FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SQ SEQUENCE 100 AA; 11675 MW; DFFCAA7561D5EFA4 CRC64;

Query Match 100.0%; Score 80; DB 1; Length 100;
Best Local Similarity 100.0%; Pred. No. 7.3e-06; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 0;

QY 1 KEGDKFKLNKSELKEL 16
Db 21 KEGDKFKLNKSELKEL 36

RESULT 2
S104 HUMAN STANDARD; PRT; 101 AA.
AC P26447;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Placental calcium-binding protein (Calvasculin) (S100 calcium-binding
DE protein A4) (MTS1 protein).
GN S100A4 OR CAPL OR MTS1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=93041710; PubMed=1384693;
RA Engelkamp D., Schaefer B.W., Erne P., Heizmann C.W.;
RT "S100 alpha, CAPL, and CACY: molecular cloning and expression
RT analysis of three calcium-binding proteins from human heart.";
RL Biochemistry 31:10258-10264(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93342029; PubMed=8341667;
RA Engelkamp D., Schaefer B., Mattei M.-G., Erne P., Heizmann C.W.;
RT "Six S100 genes are clustered on human chromosome 1q21:
RT identification of two genes coding for the two previously unreported
RT calcium-binding proteins S100D and S100E.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:6547-6551(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=93028421; PubMed=1329089;
RA Tulchinsky E.M., Ford H.L., Kramerov D., Reshetnyak E., Grigorian M.,
RA Zain S., Lukanidin E.;
RT "Transcriptional analysis of the mts1 gene with specific reference to
RT 5' flanking sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:9146-9150(1992).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix, and Prostate;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Raheij J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

```

"Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903(2002).
-!- TISSUE SPECIFICITY: Ubiquitously expressed.
-!- SIMILARITY: Belongs to the S-100 family.
-!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

DR EMBL; M80563; AAA51920.1; -;
DR EMBL; Z18950; CAA79474.1; -;
DR EMBL; Z33457; CAA83880.1; -;
DR EMBL; BC000838; AAH00838.1; -;
DR EMBL; BC016300; AAH16300.1; -;
DR PIR; A48219; A48219
DR FDB; IM31; 30-OCT-02.
DR SWISS-2DPAGE; P26447; HUMAN.
DR Genew; HGNC:10494; S100A4.
DR MIM; 114210; -;
DR GO; GO:0005509; F:calcium ion binding; TAS.
DR InterPro; IPR001751; CAPP S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CAPP S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100_CAPP; 1.
KW Calcium-binding; 3D-structure.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SQ SEQUENCE 101 AA; 11728 MW; 286D2B7B07EDB562 CRC64;
Query Match 100.0%; Score 80; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 7.4e-06; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 0;
QY 1 KEGDKFKLNKSELKEL 16
Db 22 KEGDKFKLNKSELKEL 37

RESULT 3
S104 MOUSE STANDARD; PRT; 101 AA.
AC P07051; P20066;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Placental calcium-binding protein (18A2) (PEL98) (MTS1 protein)
DE (Metastatic cell protein).
GN S100A4 OR CAPL OR MTS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=87316927; PubMed=3628004;
RA Jackson-Grusby L.L., Swiergiel J., Linzer D.I.H.;
RT "A growth-related mRNA in cultured mouse cells encodes a placental
RT calcium binding protein.";
RL Nucleic Acids Res. 15:6677-6690(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;

RX MEDLINE=88198109; PubMed=3162911;
 RA Goto K., Endo H., Fujiyoshi T.;
 RT "Cloning of the sequences expressed abundantly in established cell
 lines: identification of a cDNA clone highly homologous to S-100, a
 calcium binding protein.";
 RL J. Biochem. 103:48-53(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=89378739; PubMed=2550322;
 RX Ebralidze A., Tulchinsky E., Grigorian M., Afanasyeva A., Senin V.,
 RA Revazova E., Lukanidin E.;
 RT "Isolation and characterization of a gene specifically expressed in
 different metastatic cells and whose deduced gene product has a high
 degree of homology to a Ca2+-binding protein family.";
 RL Genes Dev. 3:1086-1093(1989).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=90236313; PubMed=2332170;
 RX Tulchinsky E.M., Grigorian M.S., Ebralidze A.K., Milshina N.I.,
 RA Lukanidin E.M.;
 RT "Structure of gene mts1, transcribed in metastatic mouse tumor
 cells.";
 RL Gene 87:219-223(1990).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93141279; PubMed=8423998;
 RX Tulchinsky E., Kramerov D., Ford H.L., Reshetnyak E., Lukanidin E.,
 RA Zain S.;
 RT "Characterization of a positive regulatory element in the mts1 gene.";
 RL Oncogene 8:79-86(1993).
 CC -!- TISSUE SPECIFICITY: Specifically expressed in different metastatic
 CC cells.
 CC -!- INDUCTION: The mRNA coding for this protein increases in
 CC abundance after serum stimulation of quiescent mouse fibroblasts.
 CC -!- SIMILARITY: Belongs to the S-100 family.
 CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X05835; CAA29282.1; -;
 DR EMBL; D00208; BAA00148.1; -;
 DR EMBL; M36578; AAA39749.1; -;
 DR EMBL; M36579; AAA39750.1; -;
 DR EMBL; X16190; CAA34316.1; -;
 DR EMBL; X16094; CAA34224.1; -;
 DR PIR; S06207; S06207.
 DR HSSP; P30801; 1A03.
 DR MGD; MGI:1330282; S100a4.
 DR InterPro; IPR001751; CaBP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF01023; S_100; 1.
 DR ProDom; PD003407; CaBP_S100; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CaBP; 1.
 DR Calcium-binding; Placenta.
 DR CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
 DR CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
 DR CONFLICT 47 47 G -> GVSXGXFNGQ (IN REF. 5).
 SQ SEQUENCE 101 AA; 11721 MW; 2302254967A4C873 CRC64;
 Query Match 96.2%; Score 77; DB 1; Length 101;
 Best Local Similarity 93.8%; Pred. No. 2.2e-05;
 Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KEGDKFKLNKSELKEL 16

Db 22 KEGDKFKLNKSELKEL 37
 RESULT 4
 S104_RAT
 ID S104_RAT STANDARD; PRT; 101 AA.
 AC P05942;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-NOV-1988 (Rel. 09, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Placental calcium-binding protein (Nerve growth factor induced protein
 DE 42A) (P9K).
 GN S100A4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88125019; PubMed=3422491;
 RA Maslakowski P., Shooter E.M.;
 RT "Nerve growth factor induces the genes for two proteins related to a
 RT family of calcium-binding proteins in PC12 cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:1277-1281(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88118907; PubMed=3430604;
 RA Barracough R., Savin J., Dube S.K., Rudland P.S.;
 RT "Molecular cloning and sequence of the gene for p9Ka. A cultured
 RT myoepithelial cell protein with strong homology to S-100, a calcium-
 RT binding protein.";
 RL J. Mol. Biol. 198:13-20(1987).
 RN [3]
 RP SEQUENCE OF 3-101 FROM N.A.
 RX MEDLINE=92158347; PubMed=1741158;
 RA de Vonge M.W., Mukherjee B.B.;
 RT "Transformation of normal rat kidney cells by v-K-ras enhances
 RT expression of transin 2 and an S-100-related calcium-binding
 RT protein.";
 RL Oncogene 7:109-119(1992).
 CC -!- INDUCTION: By nerve growth factor.
 CC -!- SIMILARITY: Belongs to the S-100 family.
 CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X06916; CAA30014.1; -;
 DR EMBL; J03628; AAA42098.1; -;
 DR EMBL; X64022; -; NOT_ANNOTATED_CDS.
 DR EMBL; X64023; -; NOT_ANNOTATED_CDS.
 DR PIR; S01759; S01759.
 DR HSSP; P30801; 1A03.
 DR InterPro; IPR001751; CaBP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF01023; S_100; 1.
 DR ProDom; PD003407; CaBP_S100; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CaBP; 1.
 DR Calcium-binding.
 DR CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
 DR CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
 SQ SEQUENCE 101 AA; 11776 MW; EA0619CEE4F487C1 CRC64;
 Query Match 90.0%; Score 72; DB 1; Length 101;
 Best Local Similarity 93.3%; Pred. No. 0.00013;

Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY	2 EGDGFKLNKSELKEL 16 : : : :
Db	23 EGDGFKLNKSELKEL 37 : : : :
RESULT 5	
S10A_BOVIN	STANDARD; PRT; 93 AA.
ID	P02639;
DT	21-JUL-1986 (Rel. 01, Created)
DT	01-APR-1988 (Rel. 07, Last sequence update)
DT	15-MAR-2004 (Rel. 43, Last annotation update)
DE	S-100 protein, alpha chain.
GN	S100A1.
OS	Bos taurus (Bovine).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC	Bovidae; Bovinae; Bos.
OX	NCBI_TaxID=9913;
RN	[1]_SEQUENCE FROM N.A.
RP	MEDLINE=86248083; PubMed=3755105;
RX	Kuwano R., Maeda T., Usui H., Araki K., Yamakuni T., Ohshima Y.,
RA	Kurihara T., Takahashi Y.;
RA	"Molecular cloning of cDNA of S100 alpha subunit mRNA.";
RT	FEBS Lett. 202:97-101(1986).
RL	[2]
RN	SEQUENCE.
RP	MEDLINE=8123652; PubMed=7250124;
RX	Isobe T., Okuyama T.;
RA	"The amino-acid sequence of the alpha subunit in bovine brain S-100a
RT	protein.";
RT	Eur. J. Biochem. 116:79-86(1981).
RL	[3]
RN	METAL ION-BINDING PROPERTIES.
RP	MEDLINE=84000339; PubMed=6615778;
RX	Baudier J., Gexard D.;
RA	"Ions binding to S100 proteins: structural changes induced by calcium
RT	and zinc on S100a and S100b proteins.";
RT	Biochemistry 22:3360-3369(1983).
CC	-!- FUNCTION: Weakly binds calcium but binds zinc very tightly-
CC	distinct binding sites with different affinities exist for both
CC	ions on each monomer. Physiological concentrations of potassium
CC	ion antagonize the binding of both divalent cations, especially
CC	affecting high-affinity calcium-binding sites.
CC	-!- SUBUNIT: Dimer of either two alpha chains, or two beta chains, or
CC	one alpha and one beta chain.
CC	-!- TISSUE SPECIFICITY: Although predominant among the water-soluble
CC	brain proteins, S-100 is also found in a variety of other tissues.
CC	-!- SIMILARITY: Belongs to the S-100 family.
CC	-!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
DR	PR; A24156; BC01A.
DR	HSP; P04631; I84C.
DR	InterPro; IPR001751; CaBP_S100.
DR	InterPro; IPR02048; EF-hand.
DR	Pfam; PF000036; ehand; 1.
DR	Pfam; PF01023; S_100; 1.
DR	ProDom; PD03407; CaBP_S100; 1.
DR	PROSITE; PS00016; EF_HAND; 1.
DR	PROSITE; PS00303; S100_CaBP; 1.
KW	Calcium-binding; zinc; Metal-binding.
FT	INIT MET 0
FT	MOD RES 1 1 BLOCKED.
FT	CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY).
FT	CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY).
FT	CONFLICT 64 64 N -> D (IN REF. 2).
SQ	SEQUENCE 93 AA; 10387 MW; 0457D0F44819B89B CRC64;
Query Match 85.0%; Score 68; DB 1; Length 93;	
Best local similarity 81.2%; Pred. No. 0.00052;	
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;	

Matches 14; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY	1 KEGDKFKLNKSELKEL 16 : : : :
Db	21 KEGDKFKLNKSELKEL 36 : : : :
RESULT 6	
S10A_HUMAN	STANDARD; PRT; 93 AA.
ID	P23297;
DT	01-NOV-1991 (Rel. 20, Created)
DT	01-NOV-1991 (Rel. 20, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	S-100 protein, alpha chain (S100 calcium-binding protein A1).
GN	S100A1 OR S100A.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]_SEQUENCE FROM N.A.
RP	TISSUE=Heart;
RX	MEDLINE=93041710; PubMed=1384693;
RA	Engelkamp D., Schaefer B.W., Erne P., Heizmann C.W.;
RT	"S100 alpha, CAPL, and CACY: molecular cloning and expression
RT	analysis of three calcium-binding proteins from human heart.";
RL	Biochemistry 31:10258-10264(1992).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Skin;
RX	MEDLINE=22388257; PubMed=12477932;
RA	Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G.,
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Bhat N.K.,
RA	Altschul S.F., Zebberg B., Buetow K.H., Schaefer C.F., Hsieh F.,
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA	Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA	Richards S., Worley K.C., Hale S., Garcia A.W., Gay L.J., Hulyk S.W.,
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA	Pahey J., Helton E., Ketting A., Madan A., Rodrigues S., Sanchez A.,
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA	Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalusz D.E.,
RA	Schmerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA	"Generation and initial analysis of more than 15,000 full-length
RT	human and mouse cDNA sequences.";
RT	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC	-!- FUNCTION: Weakly binds calcium but binds zinc very tightly-
CC	distinct binding sites with different affinities exist for both
CC	ions on each monomer. Physiological concentrations of potassium
CC	ion antagonize the binding of both divalent cations, especially
CC	affecting high-affinity calcium-binding sites.
CC	-!- SUBUNIT: Dimer of either two alpha chains, or two beta chains, or
CC	one alpha and one beta chain.
CC	-!- SUBCELLULAR LOCATION: Cytoplasmic.
CC	-!- TISSUE SPECIFICITY: Highly prevalent in heart. Also found in
CC	lesser quantities in skeletal muscle and brain.
CC	-!- SIMILARITY: Belongs to the S-100 family.
CC	-!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
This SWISS-PROT entry is copyright. It is produced through a collaboration	
between the Swiss Institute of Bioinformatics and the EMBL outstation -	
the European Bioinformatics Institute. There are no restrictions on its	
use by non-profit institutions as long as its content is in no way	
modified and this statement is not removed. Usage by and for commercial	
entities requires a license agreement (See http://www.isb-sib.ch/announce/	
or send an email to license@isb-sib.ch)	

EMBL; X58079; CAA41107.1; -.	

```

DR EMBL; BC014392; AAH14392.1; --
DR PIR; A44470; BCHUIA.
DR HSP; P04631; 1B4C.
DR Genew; HGNC:10486; S100A1.
DR MW; 176940; --
DR GO; GO:0005509; F:calcium ion binding; NAS.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0008270; F:zinc ion binding; NAS.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100 CABP; 1.
KW Calcium-binding; Zinc; Metal-binding.
FT INIT MET 0 BY SIMILARITY.
FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY).
FT CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY).
SQ SEQUENCE 93 AA; 10415 MW; 0457C356C222B89B CRC64;

Query Match 85.0%; Score 68; DB 1; Length 93;
Best Local Similarity 81.2%; Pred. No. 0.00052;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KEGDKFKLNKSELKEL 16
Db 21 KEGDKYKLSKKELKEL 36

RESULT 7
S10A_RAT
ID S10A_RAT STANDARD; PRT; 93 AA.
AC P35457;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE S-100 protein, alpha chain.
GN S100A1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP STRAIN=Sprague-Dawley;
RA Song W.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 10-93 FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=92076235; PubMed=1742602;
RA Zimmer D.B., Song W., Zimmer W.E.;
RT "Isolation of a rat S100 alpha cDNA and distribution of its mRNA in
rat tissues."
RL Brain Res. Bull. 27:157-162(1991).
CC -!- FUNCTION: Weakly binds calcium but binds zinc very tightly-
CC distinct binding sites with different affinities exist for both
CC ions on each monomer. Physiological concentrations of potassium
CC ion antagonize the binding of both divalent cations, especially
CC affecting high-affinity calcium-binding sites.
CC -!- SUBUNIT: Dimer of either two alpha chains, or two beta chains, or
CC one alpha and one beta chain.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Although predominant among the water-soluble
CC brain proteins, S-100 is also found in a variety of other tissues.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entitles requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; U26358; AAB53657.1; --
DR EMBL; U26357; AAB53657.1; JOINED.
DR EMBL; S68809; AAB20539.2; --
DR PDB; 1K2H; 13-FEB-02.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR PROSITE; PS00018; EF-hand; 1.
DR PROSITE; PS00303; S100 CABP; 1.
KW Calcium-binding; Zinc; Metal-binding; 3D-structure.
FT INIT MET 0 BY SIMILARITY.
FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY).
FT CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY).
FT CONFLICT 13 13 N -> H (IN REF. 2).
FT CONFLICT 56 56 K -> R (IN REF. 2).
SQ SEQUENCE 93 AA; 10429 MW; D2959A95EC0651A9 CRC64;

Query Match 81.2%; Score 65; DB 1; Length 93;
Best Local Similarity 75.0%; Pred. No. 0.0015;
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KEGDKFKLNKSELKEL 16
Db 21 KEGDKYKLSKKELKEL 36

RESULT 8
S102_BOVIN
ID S102_BOVIN STANDARD; PRT; 97 AA.
AC P10462;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE S100 calcium-binding protein A2 (S-100L protein).
GN S100A2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Lung;
RX MEDLINE=89139574; PubMed=2521861;
RA Glenney J.R. Jr., Kindy M.S., Zokas L.;
RT "Isolation of a new member of the S100 protein family: amino acid
sequence, tissue, and subcellular distribution."
RL J. Cell Biol. 108:569-578(1989).
CC -!- SUBUNIT: Homomultimeric (Probable).
CC -!- MISCELLANEOUS: This protein binds two calcium ions.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
DR PIR; A30129; A30129.
DR HSP; P30801; 1A03.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR PROSITE; PS00018; EF-hand; 1.
DR PROSITE; PS00303; S100 CABP; 1.
KW Calcium-binding.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SQ SEQUENCE 97 AA; 10893 MW; 55ACC9F60CF9C9E6 CRC64;

```


RL Eur. J. Biochem. 89:379-388(1978).
 RN [2]
 RP REVISIONS TO 1-4.
 RX MEDLINE=81236562; PubMed=7250124;
 RA isobe T., Okuyama T.;
 RT "The amino-acid sequence of the alpha subunit in bovine brain S-100a
 RL protein.";
 RL Eur. J. Biochem. 116:79-86(1981).
 RN [3]
 RP SEQUENCE.
 RX MEDLINE=85278169; PubMed=4026304;
 RA Marshak D.R., Umekawa H., Watterson D.M., Hidaka H.;
 RL "Structural characterization of the calcium binding protein S100 from
 RT adipose tissue.";
 RL Arch. Biochem. Biophys. 240:777-780(1985).
 RN [4]
 RP METAL ION-BINDING PROPERTIES.
 RX MEDLINE=84000339; PubMed=6615778;
 RA Baudier J., Gerard D.;
 RL "Ions binding to S100 proteins: structural changes induced by calcium
 RT and zinc on S100a and S100b proteins.";
 RL Biochemistry 22:3360-3369(1985).
 RN [5]
 RP CADMIUM-BINDING STUDIES.
 RX MEDLINE=91248136; PubMed=2039467;
 RA Donato H. Jr., Mani R.S., Kay C.M.;
 RL "Spectral studies on the cadmium-ion-binding properties of bovine
 RT brain S-100b protein.";
 RL Biochem. J. 276:113-118(1991).
 RN [6]
 RP STRUCTURE BY NMR.
 RX MEDLINE=96398693; PubMed=8805590;
 RA Kilby P.M., van Eldik L.J., Roberts G.C.K.;
 RL "The solution structure of the bovine S100b protein dimer in the
 RT calcium-free state.";
 RL Structure 4:1041-1052(1996).
 CC -1- FUNCTION: Weakly binds calcium but binds zinc very tightly-
 CC distinct binding sites with different affinities exist for both
 CC ions on each monomer. Physiological concentrations of potassium
 CC ion antagonize the binding of both divalent cations, especially
 CC affecting high-affinity calcium-binding sites.
 CC -1- SUBUNIT: Dimer of either two alpha chains, or two beta chains, or
 CC one alpha and one beta chain.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- TISSUE SPECIFICITY: Although predominant among the water-soluble
 CC brain proteins S-100 is also found in a variety of other tissues.
 CC -1- SIMILARITY: Belongs to the S-100 family.
 CC -1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 DR PIR; A91254; BCOIB.
 DR PDB; 1CFP; 12-MAR-97.
 DR PDB; 1MHO; 18-NOV-98.
 DR GO; GO:0005737; Cytoplasm; IDA.
 DR GO; GO:000576; Extracellular; ISS.
 DR GO; GO:0005509; F:calcium ion binding; ISS.
 DR GO; GO:0019210; F:kinase inhibitor activity; NAS.
 DR GO; GO:0042803; F:protein homodimerization activity; ISS.
 DR GO; GO:0048155; F:S100 alpha binding; ISS.
 DR GO; GO:0048154; F:S100 beta binding; ISS.
 DR GO; GO:0048156; F:tau protein binding; IPI.
 DR GO; GO:0008270; F:zinc ion binding; IDA.
 DR GO; GO:0048143; P:astrocyte activation; NAS.
 DR GO; GO:0007409; P:axonogenesis; NAS.
 DR GO; GO:0006112; P:energy reserve metabolism; ISS.
 DR GO; GO:0048151; P:hyperphosphorylation; NAS.
 DR GO; GO:0006917; P:induction of apoptosis; ISS.
 DR GO; GO:0007611; P:learning and/or memory; ISS.
 DR GO; GO:0045911; P:positive regulation of complement activation; NAS.
 DR GO; GO:0042035; P:regulation of cytokine biosynthesis; ISS.
 DR GO; GO:0048169; P:regulation of long-term neuronal synaptic p...; ISS.
 DR GO; GO:0006417; P:regulation of protein biosynthesis; NAS.
 DR InterPro; IPR001751; CaBP_S100.
 DR InterPro; IPR002048; EF-hand.

DR Pfam; PF00036; efhand; 1.
 DR Pfam; PF01023; S_100; 1.
 DR PROSITE; PDO03407; CaBP_S100; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CaBP; 1.
 KW Calcium-binding; Zinc; Metal-binding; Acetylation; 3D-structure.
 FT MOD_RES 1 1 ACETYLATION.
 FT CA_BIND 18 31 EF-HAND 1 (LOW AFFINITY).
 FT CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY).
 FT HELIX 2 18
 FT TURN 19 19
 FT TURN 24 25
 FT STRAND 27 27
 FT HELIX 29 39
 FT TURN 41 43
 FT HELIX 50 60
 FT STRAND 68 68
 FT HELIX 70 87
 SQ SEQUENCE 91 AA; 10537 MW; 386201933DE6B93A CRC64;
 Query Match 80.0%; Score 64; DB 1; Length 91;
 Best Local Similarity 81.2%; Pred. No. 0.0022;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 KEGDKFKLNKSELKEL 16
 Db 20 REGDKHKLKSELKEL 35
 RESULT 11
 S10B_HUMAN
 ID S10B_HUMAN STANDARD; PRT; 91 AA.
 AC P04271;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 20-MAR-1987 (Rel. 04, Last sequence update)
 DE 10-OCT-2003 (Rel. 42, Last annotation update)
 DE S-100 protein, beta chain.
 GN S100B.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=85291729; PubMed=4031854;
 RA Jensen R., Marshak D.R., Anderson C., Lukas T.J., Watterson D.M.;
 RT "Characterization of human brain S100 protein fraction: amino acid
 RT sequence of S100 beta.";
 RL J. Neurochem. 45:700-705(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90368757; PubMed=2394738;
 RA Allore R.J., Friend W.C., O'Hanlon D., Neilson K.M., Bauml R.,
 RA Dunn R.J., Marks A.;
 RT "Cloning and expression of the human S100 beta gene.";
 RL J. Biol. Chem. 265:15537-15543(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Skallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 [4]
 RP METAL ION-BINDING PROPERTIES.
 RX MEDLINE=85023393; PubMed=6487634;
 RA Baudier J., Glasser N., Haglid K., Gerard D.,
 RT "Purification, characterization and ion binding properties of human
 RL brain S100b protein.";
 RL Biochim. Biophys. Acta 790:164-173 (1984).
 [5]
 RP STRUCTURE BY NMR.
 RX MEDLINE=98179937; PubMed=9519411;
 RA Smith S.P., Shaw G.S.,
 RT "A novel calcium-sensitive switch revealed by the structure of human
 RL S100B in the calcium-bound form.";
 RL Structure 6:211-222 (1998).
 CC -!- FUNCTION: Weakly binds calcium but binds zinc very tightly-
 CC ions on each monomer. Physiological concentrations of potassium
 CC ion antagonize the binding of both divalent cations, especially
 CC affecting high-affinity calcium-binding sites.
 CC -!- SUBUNIT: Dimer of either two alpha chains, or two beta chains, or
 CC one alpha and one beta chain.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- TISSUE SPECIFICITY: Although predominant among the water-soluble
 CC brain proteins, S-100 is also found in a variety of other tissues.
 CC -!- MISCELLANEOUS: In addition to metal-ion binding, this protein is
 CC involved with the regulation of protein phosphorylation in brain
 CC tissue.
 CC -!- SIMILARITY: Belongs to the S-100 family.
 CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 DR EMBL; M59488; AAA60367.1; -;
 DR EMBL; M59487; AAA60367.1; JOINED.
 DR EMBL; BC001766; AAH01766.1; -;
 DR FIR; A38364; BCHUB.
 DR PDB; 1UWO; 10-JUN-98.
 DR PDB; 1WQ1; 25-DEC-02.
 DR Genew; HGNC:10500; S100B.
 DR MIM; 176990; -;
 DR GO; GO:005737; C:cytoplasm; ISS.
 DR GO; GO:0005576; C:extracellular; ISS.
 DR GO; GO:0005509; F:calcium ion binding; NAS.
 DR GO; GO:0019210; F:kinase inhibitor activity; ISS.
 DR GO; GO:0042803; F:protein homodimerization activity; IPI.
 DR GO; GO:0048153; F:S100 alpha binding; IPI.
 DR GO; GO:0048154; F:S100 beta binding; IPI.
 DR GO; GO:0048156; F:tau protein binding; ISS.
 DR GO; GO:0008270; F:zinc ion binding; ISS.
 DR GO; GO:0048143; F:astrocyte activation; ISS.
 DR GO; GO:0007409; P:axonogenesis; ISS.
 DR GO; GO:0008283; P:calcium ion homeostasis; ISS.
 DR GO; GO:0007417; P:cell proliferation; TAS.
 DR GO; GO:0006112; P:energy reserve metabolism; ISS.
 DR GO; GO:0048151; P:hyperphosphorylation; ISS.
 DR GO; GO:0006917; P:induction of apoptosis; ISS.
 DR GO; GO:0007611; P:learning and/or memory; ISS.
 DR GO; GO:0045917; P:positive regulation of complement activation; ISS.
 DR GO; GO:0042035; P:regulation of cytokine biosynthesis; ISS.
 GO; GO:0048169; P:regulation of long-term neuronal synaptic p...; ISS.
 GO; GO:0006417; P:regulation of protein biosynthesis; ISS.
 InterPro; IPR001751; CaBP_S100.
 InterPro; IPR002048; EF-hand.
 Pfam; PF00036; ehand; 1.
 Pfam; PF01023; S100; 1.
 ProDom; PD003407; CaBP_S100; 1.
 ProSITE; PS00018; EF_HAND; 1.
 ProSITE; PS00303; S100_CASP; 1.
 Calcium-binding; Zinc; Metal-binding; 3D-structure.
 INIT MET 0 0
 FT MOD RES 1 1 BLOCKED.
 FT CA_BIND 18 31 EF-HAND 1 (LOW AFFINITY).
 FT CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY).
 FT HELIX 2 20
 FT STRAND 28 28
 FT TURN 29 29
 FT HELIX 30 38
 FT TURN 39 39
 FT HELIX 52 60
 FT STRAND 67 67
 FT HELIX 70 86
 FT TURN 87 87
 SQ SEQUENCE 91 AA; 10582 MW; 2378AA8B8FEA79F6 CRC64;
 Query Match 80.0%; Score 64; DB 1; Length 91;
 Best Local Similarity 81.2%; Pred. No. 0.0022;
 Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 KEGDKFKLNKSELKEL 16
 :|||||
 Db 20 REGDKHKLKSELKEL 35
 RESULT 12
 S100B MOUSE
 ID S10B MOUSE STANDARD; PRT; 91 AA.
 AC P50114;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE S-100 protein, beta chain.
 GN S100B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c;
 RX MEDLINE=93388628; PubMed=8376406;
 RA Jiang H., Shah S., Hilt D.C.;
 RT "Organization, sequence, and expression of the murine S100 beta gene.
 RT Transcriptional regulation by cell type-specific cis-acting
 RT regulatory elements.";
 RL J. Biol. Chem. 268:20502-20511 (1993).
 CC -!- FUNCTION: Weakly binds calcium but binds zinc very tightly-
 CC distinct binding sites with different affinities exist for both
 CC ions on each monomer. Physiological concentrations of potassium
 CC ion antagonize the binding of both divalent cations, especially
 CC affecting high-affinity calcium-binding sites.
 CC -!- SUBUNIT: Dimer of either two alpha chains, or two beta chains, or
 CC one alpha and one beta chain.
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- TISSUE SPECIFICITY: Although predominant among the water-soluble
 CC brain proteins, S-100 is also found in a variety of other tissues
 CC (by similarity).
 CC -!- SIMILARITY: Belongs to the S-100 family.
 CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its


```

DR PDB; 1MWN; 18-DEC-02.
DR GO; GO:0005737; C:cytoplasm; ISS.
DR GO; GO:0005576; C:extracellular; ISS.
DR GO; GO:0005509; F:calcium ion binding; ISS.
DR GO; GO:0019210; F:kinase inhibitor activity; ISS.
DR GO; GO:0048156; F:tau protein binding; ISS.
DR GO; GO:0008270; F:zinc ion binding; ISS.
DR GO; GO:0048143; F:astrocyte activation; ISS.
DR GO; GO:0007409; P:axogenesis; ISS.
DR GO; GO:0006874; P:calcium ion homeostasis; ISS.
DR GO; GO:0006112; P:energy reserve metabolism; ISS.
DR GO; GO:0048151; P:hyperphosphorylation; ISS.
DR GO; GO:0006917; P:induction of apoptosis; ISS.
DR GO; GO:0007611; P:learning and/or memory; ISS.
DR GO; GO:0045917; P:positive regulation of complement activation; ISS.
DR GO; GO:0042035; P:regulation of cytokine biosynthesis; ISS.
DR GO; GO:0048169; P:regulation of long-term neuronal synaptic p. .; ISS.
DR GO; GO:0006417; P:regulation of protein biosynthesis; ISS.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CABP_S100; 1.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100 CABP; 1.
KW Calcium-binding; zinc; Metal-binding; 3D-structure.
FT INIT MET 0
FT CA_BIND 18 31 EF-HAND 1 (LOW AFFINITY).
FT CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY).
FT HELIX 2 17
FT TURN 18 19
FT STRAND 27 27
FT TURN 29 39
FT HELIX 41 42
FT HELIX 43 46
FT HELIX 50 61
FT TURN 62 62
FT STRAND 68 68
FT TURN 70 83
FT TURN 87 88
SQ SEQUENCE 91 AA; 10613 MW; 2378AA8B8F7134D CRC64;

Query Match 80.0%; Score 64; DB 1; Length 91;
Best Local Similarity 81.2%; Pred. No. 0.0022; 1; Mismatches 0; Gaps 0;
Matches 13; Conservative 1; Indels 2;

QY 1 KEGDKFKLNKSELKEL 16
:|||||:|||||
DB 20 REGDKHKLNKSELKEL 35

RESULT 14
S10A_MOUSE STANDARD; PRT; 93 AA.
AC P56565; O88949;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE S-100 protein, alpha chain (S100 calcium-binding protein A1).
GN S100A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RA Marra M., Hillier L., Allen M., Bowles M., Dietrich N., Dubuque T.,
RA Geisel S., Kucaba T., Lacy M., Le M., Martin J., Morris M.,
RA Schellenberg K., Steptoe M., Tan F., Underwood K., Moore B.,
RA Theising B., Wylie T., Lennon G., Soares B., Wilson R., Waterston R.,
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```

```

RX MEDLINE=99117144; PubMed=9920416;
RA Ridinger K., Ilg E.C., Niggli F.K., Heizmann C.W., Schaefer B.W.;
RT "Clustered organization of S100 genes in human and mouse.";
RL Biochim. Biophys. Acta 1498:254-263(1998).
CC -!- FUNCTION: Weakly binds calcium but binds zinc very tightly-
CC distinct binding sites with different affinities exist for both
CC ions on each monomer. Physiological concentrations of potassium
CC ion antagonize the binding of both divalent cations, especially
CC affecting high-affinity calcium-binding sites (By similarity).
CC -!- SUBUNIT: Dimer of either two alpha chains, or two beta chains, or
CC one alpha and one beta chain (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AA000715; -; NOT ANNOTATED_CDS.
DR EMBL; AA207749; -; NOT ANNOTATED_CDS.
DR EMBL; AA500563; -; NOT ANNOTATED_CDS.
DR EMBL; AA432539; -; NOT ANNOTATED_CDS.
DR EMBL; AF087687; AAC64108.1; -.
DR HSSP; P04631; IB4C.
DR MGD; MGI:1338917; S100a1.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR ProDom; PD003407; CABP_S100; 1.
DR PROSITE; PS00018; EF-HAND; 1.
DR PROSITE; PS00018; EF-HAND; 1.
DR PROSITE; PS00303; S100 CABP; 1.
KW Calcium-binding; zinc; Metal-binding.
FT INIT MET 0
FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY) (BY SIMILARITY).
FT CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY) (BY SIMILARITY).
FT CONFLICT 21 21 Q -> E (IN AA207749).
FT CONFLICT 21 21 Q -> M (IN AA500563).
FT CONFLICT 21 21 Q -> H (IN AA432539).
FT CONFLICT 21 21 Q -> K (IN REF. 2).
SQ SEQUENCE 93 AA; 10374 MW; C96E39EA18D705C CRC64;

Query Match 76.2%; Score 61; DB 1; Length 93;
Best Local Similarity 68.8%; Pred. No. 0.0066; 4; Mismatches 1; Indels 0; Gaps 0;
Matches 11; Conservative 4;

QY 1 KEGDKFKLNKSELKEL 16
:|||||:|||||
DB 21 QEGDKYKLNKSELKEL 36

RESULT 15
S106_MOUSE STANDARD; PRT; 89 AA.
AC P14069;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Calyculin (Protein) receptor associated protein (5B10).
GN S100A6 OR CACY.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.

```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:44:32; Search time 8.15484 Seconds
(without alignments)
188.730 Million cell updates/sec

Title: US-10-067-618-5

Perfect score: 80

Sequence: 1 KEGDKFKLNKSELKEL 16

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

PIR_78.*
1: PIR1.*
2: PIR2.*
3: PIR3.*
4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	100.0	100	2 A53217	placental calcium-
2	80	100.0	101	2 A48219	calvasculin - huma
3	77	96.2	101	2 S06207	calvasculin - mous
4	72	90.0	101	2 S01759	calvasculin - rat
5	68	85.0	94	1 BCU1A	S-100 protein alph
6	68	85.0	94	1 BCU1A	S-100 protein alph
7	65	81.2	97	2 A30129	S-100 protein, lun
8	65	81.2	98	2 A41988	S-100 calcium-bind
9	64	80.0	91	1 BCU1B	S-100 protein beta
10	64	80.0	92	1 BCU1B	S-100 protein beta
11	64	80.0	92	2 A26557	S-100 protein beta
12	64	80.0	92	2 A48015	S-100 protein beta
13	64	80.0	95	1 S35985	S-100 protein alph
14	58	72.5	99	2 A54314	calcyclin - mouse
15	58	72.5	90	1 S27011	calcyclin - rabbit
16	58	72.5	90	2 B28363	calcyclin - rat
17	55	68.8	90	1 BCU1Y	calcyclin - human
18	51	63.7	101	2 C48219	S-100 calcium-bind
19	49	61.3	79	1 KLPGT	calcium-binding pr
20	47	58.8	110	2 B48219	S-100 calcium-bind
21	46.5	58.1	461	2 B70379	cell division prote
22	46	57.5	277	2 B72226	hypothetical prote
23	45	56.2	1225	2 A49464	chromosome segrega
24	44.5	55.6	58	2 H69295	conserved hypothet
25	44	55.0	79	1 JN0246	calcium-binding pr
26	44	55.0	79	1 KLRTI	calcium-binding pr
27	44	55.0	79	1 KLBOI	calcium-binding pr
28	44	55.0	519	2 D69006	DNA-directed RNA p
29	43.5	54.4	328	2 A64471	hypothetical prote

30 43 53.8 227 2 D90512 conserved hypothet
31 43 53.8 480 2 B86689 prophage ps2 prote
32 43 53.8 1173 2 S48877 Ca2+-transporting
33 42.5 53.1 339 2 S76686 hypothetical prote
34 42 52.5 176 2 A71457 hypothetical prote
35 42 52.5 475 2 S53389 glutamate decarbox
36 42 52.5 490 2 B86265 cytochrome P450 71
37 42 52.5 621 2 A71961 90kDa chaperone -
38 42 52.5 743 2 AG2716 reductase [impor
39 42 52.5 743 2 B97498 molybdopterin oxid
40 42 52.5 853 2 G30559 trse-like protein
41 42 52.5 1031 2 I51674 gastric H(+)-K(+)-
42 41.5 51.9 208 2 D89767 hypothetical prote
43 41 51.2 65 2 A41004 calgizzarin - chic
44 41 51.2 76 2 S60789 M protein precurso
45 41 51.2 147 2 S65579 hypothetical prote

ALIGNMENTS

RESULT 1

A53217

Placental calcium-binding protein homolog - bovine

C/Species: Bos primigenius taurus (cattle)

C/Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 17-Mar-1999

C/Accession: A53217

R/Polans, A.S.; Palczewski, K.; Asson-Batres, M.A.; Ohguro, H.; Witkowska, D.; Haley, T.I.
J. Biol. Chem. 269, 6233-6240, 1994

A/Title: Purification and primary structure of capl, an S-100-related calcium-binding pr

A/Reference number: A53217; MUID:94164991; PMID:8119967

A/Accession: A53217

A/Status: preliminary

A/Molecule type: protein

A/Residues: 1-100 <POL>

C/Superfamily: S-100 protein; calmodulin repeat homology

C/Keywords: acetylated amino end; calcium binding; EF hand

F/6-40/Domain: calmodulin repeat homology <EF1>

F/1/Modified site: acetylated amino end (Ala) #status experimental

Query Match

Best Local Similarity 100.0%; Score 80; DB 2; Length 100;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KEGDKFKLNKSELKEL 16

Db 21 KEGDKFKLNKSELKEL 36

RESULT 2

A48219

calvasculin - human

N/Alternate names: calcium-binding protein L (CAPL); metastasin; placental calcium-bindir

C/Species: Homo sapiens (man)

C/Date: 02-Jun-1994 #sequence_revision 02-Jun-1994 #text_change 13-Aug-1999

C/Accession: A48219; B44470; FC1223

R/Engelkamp, D.; Schaefer, B.W.; Mattei, M.G.; Erne, P.; Heizmann, C.W.

Proc. Natl. Acad. Sci. U.S.A. 90, 6547-6551, 1993

A/Title: Six S100 genes are clustered on human chromosome 1q21: identification of two ge

A/Reference number: A48219; MUID:93342029; PMID:8341667

A/Accession: A48219

A/Status: nucleic acid sequence not shown

A/Molecule type: DNA

A/Residues: 1-101 <EN>

A/Cross-references: GB:Z18950; NID:G396708; PIDN:CAA79474.1; PID:G396710

R/Engelkamp, D.; Schaefer, B.W.; Erne, P.; Heizmann, C.W.

Biochemistry 31, 10258-10264, 1992

A/Title: S100 alpha, CAPL, and CACY: molecular cloning and expression analysis of three c

A/Reference number: A44470; MUID:93041710; PMID:1384693

A/Accession: B44470

A/Status: not compared with conceptual translation

A/Molecule type: mRNA

A/Residues: 1-101 <EN2>

A;Cross-references: GB:M80563; GB:M77499; NID:g179916; PIDN:AAAS1920.1; PID:g179917
A;Experimental source: heart
A;Note: sequence extracted from NCB1 backbone (NCBIP:116496)
R;Tomida, Y.; Terasawa, M.; Kobayashi, R.; Hidaka, H.
Biochem. Biophys. Res. Commun. 189, 1310-1316, 1992
A;Title: Calcyclin and calvasculin exist in human platelets.
A;Reference number: PC1222; MUID:93129189; PMID:1482346
A;Accession: PC1223
A;Molecule type: protein
A;Residues: 8-18;36-57 <COM>
A;Experimental source: platelets
C;Genetics:
A;Gene: GDB:S100A4; CAPL
A;Cross-references: GDB:119748; OMIM:114210
A;Map position: 1q21-1q21
A;Introns: 47/3
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: calcium binding; EF hand
F;7-41/Domain: calmodulin repeat homology <EF1>
F;50-82/Domain: calmodulin repeat homology <EF2>

Query Match 100.0%; Score 80; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 2.5e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
Db 22 KEGDKFKLNKSELKEL 37

RESULT 3
S06207
calvasculin - mouse
N;Alternate names: calcium-binding protein mts1; calcium-binding protein pEL98; placenta
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1990 #sequence revision 31-Dec-1990 #text change 20-Jun-2000
R;Accession: S06207; JH0097; S07981; A26803; A41411; I48674
R;Ebraldize, A.; Tulchinsky, E.; Grigorian, M.; Afanas'yeva, A.; Serin, V.; Revazova, E.;
Genes Dev. 3, 1086-1093, 1989
A;Title: Isolation and characterization of a gene specifically expressed in different me
A;Reference number: S06207; MUID:89378739; PMID:2550322
A;Accession: S06207
A;Molecule type: mRNA
A;Residues: 1-101 <EBR>
A;Cross-references: EMBL:X16190; NID:g54926; PIDN:CAA34316.1; PID:g54927
R;Tulchinsky, E.M.; Grigorian, M.S.; Ebraldize, A.K.; Milshina, N.I.; Lukanidin, E.M.
Gene 87, 219-223, 1990
A;Title: Structure of gene mts1, transcribed in metastatic mouse tumor cells.
A;Reference number: JH0097; MUID:90236313; PMID:2332170
A;Accession: JH0097
A;Molecule type: DNA
A;Residues: 1-101 <TUL>
A;Cross-references: GB:M36578; GB:M36579
A;Experimental source: liver
R;Tulchinsky, B.
submitted to the EMBL Data Library, August 1989
A;Reference number: S07981
A;Accession: S07981
A;Molecule type: DNA
A;Residues: 1-47, 'VSGSXFXNG', 56-57, 'RTDERA', <TUL>
A;Cross-references: EMBL:X16034; NID:g53249; PIDN:CAA34224.1; PID:g53250
R;Jackson-Grusby, L.L.; Swergel, J.; Linzer, D.I.H.
Nucleic Acids Res. 15, 6677-6690, 1987
A;Title: A growth-related mRNA in cultured mouse cells encodes a placental calcium bindi
A;Reference number: A26803; MUID:87316927; PMID:3628004
A;Accession: A26803
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-101 <JAC>
A;Cross-references: GB:X05835; NID:g50310; PIDN:CAA29282.1; PID:g50311
R;Goto, K.; Endo, H.; Fujiyoshi, T.
J. Biochem. 103, 48-53, 1988

A;Title: Cloning of the sequences expressed abundantly in established cell lines: identifi
A;Reference number: A41411; MUID:88198109; PMID:3162911
A;Accession: A41411
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-101 <GOT>
A;Cross-references: GB:D00208; NID:g220569; PIDN:BAA00148.1; PID:g220570
R;Tulchinsky, E.; Kramerov, D.; Ford, H.L.; Reshetnyak, E.; Lukanidin, E.; Zain, S.
Oncogene 8, 79-86, 1993
A;Title: Characterization of a positive regulatory element in the mts1 gene.
A;Reference number: I48674; MUID:93141279; PMID:8423998
A;Accession: I48674
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-47, 'VSGSXFXNG', 48-54 <RBS>
A;Cross-references: EMBL:X16094; NID:g53249; PIDN:CAA34224.1; PID:g53250
C;Comment: Gene mts1 is expressed in metastatic cells.
C;Genetics:
A;Gene: mts1
A;Introns: 47/3
C;Superfamily: S-100 protein; calmodulin repeat homology
C;Keywords: calcium binding; cancer; EF hand
F;7-41/Domain: calmodulin repeat homology <EF1>
F;50-82/Domain: calmodulin repeat homology <EF2>

Query Match 96.2%; Score 77; DB 2; Length 101;
Best Local Similarity 93.8%; Pred. No. 7.4e-05;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
Db 22 KEGDKFKLNKSELKEL 37

RESULT 4
S01759
calvasculin - rat
N;Alternate names: calcium-binding protein p9Ka/42A; gene p9Ka protein
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1989 #sequence revision 30-Jun-1989 #text_change 13-Aug-1999
R;Accession: S01759; B31373; S26496; S26497
R;Barraclough, R.; Savin, J.; Dube, S.K.; Rudland, P.S.
J. Mol. Biol. 198, 13-20, 1987
A;Title: Molecular cloning and sequence of the gene for p9Ka : a cultured myoepithelial
A;Reference number: S01759; MUID:88118907; PMID:3430604
A;Accession: S01759
A;Molecule type: DNA
A;Residues: 1-101 <BAR>
A;Cross-references: EMBL:X06916; NID:g56832; PIDN:CAA30014.1; PID:g56833
R;Maslakowski, P.; Shooter, E.M.
Proc. Natl. Acad. Sci. U.S.A. 85, 1277-1281, 1988
A;Title: Nerve growth factor induces the genes for two proteins related to a family of c
A;Reference number: A94189; MUID:88125019; PMID:3422491
A;Accession: B31373
A;Molecule type: mRNA
A;Residues: 1-101 <MAS>
A;Cross-references: GB:J03628; NID:g206829; PIDN:AAA42098.1; PID:g206830
R;de Vongue, M.W.; Mukherjee, B.B.
Oncogene 7, 109-119, 1992
A;Title: Transformation of normal rat kidney cells by v-K-ras enhances expression of tra
A;Reference number: S26496; MUID:92158347; PMID:1741158
A;Accession: S26496
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 65-100 <DEV>
A;Cross-references: EMBL:X64023
A;Accession: S26497
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 3-97 <DE2>
A;Cross-references: EMBL:X64022
C;Genetics:
A;Gene: p9Ka

A:Molecule type: protein
A:Residues: 2-64,'D',66-94 <ISO>
R:Baudier, J.; Gerard, D.
Biochemistry 22, 3360-3369, 1983
A:Title: Ions binding to S100 proteins: structural changes induced by calcium and zinc
A:Reference number: A90471; PMID:84000339; PMID:6615778
A:Contents: annotation; metal ion-binding properties
R:Okazaki, K.; Obata, N.H.; Inoue, S.; Hidaka, H.
Biochem. J. 306, 551-555, 1995
A:Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in glia
A:Reference number: S54343; PMID:95194333; PMID:7887910
A:Accession: S54346
A:Molecule type: protein
A:Residues: 24-33 <OKA>
C:Comment: The S-100 protein is composed of two related polypeptide chains, alpha and beta
brain proteins, S-100 is also found in a variety of other tissues.
C:Comment: S-100 is an intracellular protein that weakly binds calcium. It binds zinc ver
es, with different affinities, exist for both ions on each monomer. Physiological concent
alcium-binding sites.
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: blocked amino end; brain; calcium binding; EF hand; zinc
F:2-94/Product: S-100 protein alpha chain #status predicted <MAT>
F:7-41/Domain: calmodulin repeat homology <EF1>
F:50-82/Domain: calmodulin repeat homology <EF2>
F:2/Modified site: blocked amino end (Gly) (in mature form) (probably acetylated) #status
F:20,23,25,28,33/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted
F:63,65,67,69,74/Binding site: calcium (Asp, Asn, Asp, Glu, Glu) #status predicted

Query Match 85.0%; Score 68; DB 1; Length 94;
Best Local Similarity 81.2%; Pred. No. 0.0017;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
|||||:|:|:|:|
DB 22 KEGDKYKLSKKELKEL 37

RESULT 7
A30129
S-100 protein, lung - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 01-Sep-1995
C:Accession: A30129
R:Glenney Jr., J.R.; Kindy, M.S.; Zokas, L.
J. Cell Biol. 108, 569-578, 1989
A:Title: Isolation of a new member of the S100 protein family: amino acid sequence, tiss
A:Reference number: A30129; PMID:89139574; PMID:2521861
A:Accession: A30129
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-97 <GLE>
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; EF hand
F:50-82/Domain: calmodulin repeat homology <EF2>

Query Match 81.2%; Score 65; DB 2; Length 97;
Best Local Similarity 75.0%; Pred. No. 0.005;
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
:|||||:|:|:|
DB 22 QEGDKPKLSKGENKEL 37

RESULT 8
A41988
S-100 calcium-binding protein A2 - human
N:Alternate names: calcium-binding protein Can19; S-100 calcium-binding protein L
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 29-Aug-1997
C:Accession: A41988
R:Lee, S.W.; Tomasetto, C.; Swisshelm, K.; Keyomarsi, K.; Sager, R.
Proc. Natl. Acad. Sci. U.S.A. 89, 2504-2508, 1992

A>Title: Down-regulation of a member of the S100 gene family in mammary carcinoma cells
A:Reference number: A41988; MUID:92196147; PMID:1372446
A:Accession: A41988
A:Molecule type: mRNA
A:Residues: 1-98 <LEE>
A:Cross-references: GB:M87068; NID:g179896
A:Experimental source: mammary carcinoma cells
A:Note: Sequence extracted from NCBI backbone (NCBIN:88561, NCBIP:88562)
C:Genetics:
A:Gene: GDB:S100A2; CMM19; S100L
A:Cross-references: GDB:211118; OMIM:176993
A:Map position: 1q21-1q21
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; EF hand
F:8-42/Domain: calmodulin repeat homology <EF1>
F:51-83/Domain: calmodulin repeat homology <EF2>

Query Match 81.2%; Score 65; DB 2; Length 98;
Best Local Similarity 75.0%; Pred. No. 0.005;
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
:|||||:|:|
DB 23 QEGDKFKLSKGMKEL 38

RESULT 9
BCBOIB
S-100 protein beta chain - bovine
N:Alternate names: neurocalcin delta-binding protein S100-beta
C:Species: Bos primigenius taurus (cattle)
C:Date: 31-May-1979 #sequence revision 14-Nov-1983 #text_change 24-Nov-1999
C:Accession: A91254; B91110; A90075; S54348; A03077
R:Isobe, T.; Okuyama, T.
J. Biol. Chem. 89, 379-388, 1978
A>Title: The amino-acid sequence of S-100 protein (PAP-I-b protein) and its relation to
A:Reference number: A91254; MUID:79045265; PMID:710399
A:Accession: A91254
A:Molecule type: protein
A:Residues: 'ESEL', 5-91 <ISO>
A:Note: This sequence has since been revised in reference A91110
R:Isobe, T.; Okuyama, T.
Eur. J. Biochem. 116, 79-86, 1981
A>Title: The amino-acid sequence of the alpha subunit in bovine brain S-100 a protein.
A:Reference number: A91110; MUID:81236562; PMID:7250124
A:Accession: B91110
A:Molecule type: protein
A:Residues: 1-91 <IS2>
R:Baudier, J.; Gerard, D.
Biochemistry 22, 3360-3369, 1983
A>Title: Ions binding to S100 proteins: structural changes induced by calcium and zinc
A:Reference number: A90471; MUID:84000339; PMID:6615778
A:Contents: annotation; metal ion-binding properties
R:Marshak, D.R.; Umekawa, H.; Watterson, D.M.; Hidaka, H.
Arch. Biochem. Biophys. 240, 777-780, 1985
A>Title: Structural characterization of the calcium binding protein S100 from adipose ti
A:Reference number: A90075; MUID:85278169; PMID:4026304
A:Accession: A90075
A:Molecule type: protein
A:Residues: 1-91 <MAR>
R:Okazaki, K.; Obata, N.H.; Inoue, S.; Hidaka, H.
Biochem. J. 306, 551-555, 1995
A>Title: S100-beta is a target protein of neurocalcin delta, an abundant isoform in glia
A:Reference number: S54343; MUID:95194333; PMID:7887910
A:Accession: S54348
A:Molecule type: protein
A:Residues: 56-61, 'N', 63-79, 'V' <OKA>
C:Comment: The S-100 protein is composed of two related polypeptide chains, alpha and be
brain proteins, S-100 is also found in a variety of other tissues.
C:Comment: S-100 is an intracellular protein that weakly binds calcium. It binds zinc ve
es with different affinities exist for both ions on each monomer. Physiological concentr

cium-binding sites.
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: blocked amino end; brain; calcium binding; EF hand; zinc
F:5-39/Domain: calmodulin repeat homology <EF1>
F:48-80/Domain: calmodulin repeat homology <EF2>
F:1/Modified site: blocked amino end (Ser) (probably acetylated) #status experimental
F:18-21,23,26,31/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted
F:61,63,65,67,72/Binding site: calcium (Asp, Asp, Glu, Glu) #status predicted

Query Match 80.0%; Score 64; DB 1; Length 91;
Best Local Similarity 81.2%; Pred. No. 0.0067;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
:|||||:|:|
DB 20 REGDKHKLNKSELKEL 35

RESULT 10
BCHUIB
S-100 protein beta chain [validated] - human
N:Alternate names: neural S-100 calcium-binding protein beta
C:Species: Homo sapiens (man)
C:Date: 04-Dec-1986 #sequence revision 06-Jan-1995 #text_change 08-Dec-2000
C:Accession: A38364; A92972; A03076
R:Allore, R.J.; Friend, W.C.; O'Hanlon, D.; Neilson, K.M.; Baunmal, R.; Dunn, R.J.; Marks
J. Biol. Chem. 265, 15537-15543, 1990
A>Title: Cloning and expression of the human S100beta gene.
A:Reference number: A38364; MUID:90368757; PMID:2394738
A:Accession: A38364
A:Molecule type: DNA
A:Residues: 1-92 <ALL>
A:Cross-references: GB:J05600; GB:M59486; NID:g337726; GB:M59487; NID:g337727; GB:M59488
R:Jensen, R.; Marshak, D.R.; Anderson, C.; Lukas, T.O.; Watterson, D.M.
J. Neurochem. 45, 700-705, 1985
A>Title: Characterization of human brain S100 protein fraction: amino acid sequence of S
A:Reference number: A92972; MUID:85291729; PMID:4031854
A:Accession: A92972
A:Molecule type: protein
A:Residues: 2-92 <JEN>
R:Baudier, J.; Glasser, N.; Haglid, K.; Gerard, D.
Biochim. Biophys. Acta 790, 164-173, 1984
A>Title: Purification, characterization and ion binding properties of human brain S100b
A:Reference number: A90653; MUID:85023393; PMID:6487634
A:Contents: annotation; metal ion-binding properties
C:Comment: this protein binds p53, tubulin and many other proteins at physiological conc
C:Comment: S-100 is an intracellular protein that binds calcium. It binds zinc more tigh
different affinities exist for both ions on each monomer. Physiological concentrat
nding sites.
C:Comment: This protein is expressed predominantly in brain tissue by astroglial cells.
C:Comment: The homodimer contains disulfide bonds, but the bond pattern has not been det
C:Genetics:
A:Gene: GDB:S100B
A:Cross-references: GDB:120360; OMIM:176990
A:Map position: 21q22.3-21q22.3
A:Introns: 46/3
A:Note: the first intron occurs before the initiator codon
C:Complex: homodimer; heterodimer with S-100 protein alpha chain (see PIR:BCHUIA)
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: blocked amino end; brain; calcium binding; EF hand; heterodimer; homodimer;
F:2-92/Product: S-100 protein beta chain #status experimental <MAR>
F:6-40/Domain: calmodulin repeat homology <EF1>
F:49-81/Domain: calmodulin repeat homology <EF2>
F:2/Modified site: blocked amino end (Ser) (in mature form) (probably acetylated) #statu
F:19,22,24,27,32/Binding site: calcium (Ser, Glu, Asp, Lys, Glu) #status predicted
F:62,64,66,68,73/Binding site: calcium (Asp, Asp, Glu, Glu) #status predicted

Query Match 80.0%; Score 64; DB 1; Length 92;
Best Local Similarity 81.2%; Pred. No. 0.0067;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
:|||||:|:|

Db 21 REGDKHKLKSELKEL 36

RESULT 11

A26557
S-100 protein beta chain - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 25-Mar-1988 #sequence_revision 04-Nov-1994 #text_change 13-Aug-1999
C:Accession: A60046; S07357; A26557
R:Kuwano, R.; Usui, H.; Araki, K.; Kuwano, R.; Takahashi, Y.; Suzuki, Y.
Brain Res. Mol. Brain Res. 10, 193-202, 1991
A:Title: Structure and expression of rat S-100 beta subunit gene.
A:Reference number: A60046; MUID:91359841; PMID:1653388
A:Accession: A60046
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-92 <MAE>
A:Cross-references: GB:S53527
R:Kuwano, R.; Usui, H.; Maeda, T.; Fukui, T.; Yamanari, N.; Ohtsuka, E.; Ikehara, M.; Tanabe, T.; Usui, H.; Araki, K.; Kuwano, R.; Takahashi, Y.; Suzuki, Y.
Nucleic Acids Res. 12, 7455-7465, 1984
A:Title: Molecular cloning and the complete nucleotide sequence of cDNA to mRNA for S-100 beta subunit gene.
A:Reference number: S07357; MUID:85037924; PMID:6093041
A:Accession: S07357
A:Molecule type: mRNA
A:Residues: 1-92 <KUW>
A:Cross-references: EMBL:X01090; NID:g57174; PIDN:CAA25567.1; PID:g57175
R:Dunn, R.; Landry, C.; O'Hanlon, D.; Dunn, J.; Allaire, R.; Brown, I.; Marks, A.
J. Biol. Chem. 262, 3562-3566, 1987
A:Title: Reduction in S100 protein beta-subunit mRNA in C6 rat glioma cells following treatment with retinoic acid.
A:Reference number: A26557; MUID:87137648; PMID:3818655
A:Accession: A26557
A:Molecule type: mRNA
A:Residues: 6-92 <DUN>
A:Cross-references: GB:M15705
C:Comment: S-100 protein occurs as alpha-beta heterodimers, alpha-alpha heterodimers, and alpha-gamma heterodimers.
C:Genetics:
A:Introns: 46/3
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: brain; calcium binding; dimer; EF hand; zinc
F:2-92/Product: S-100 protein beta chain #status predicted <MAT>
F:6-40/Domain: calmodulin repeat homology <EF1>

Query Match 80.0%; Score 64; DB 2; Length 92;

Best Local Similarity 81.2%; Pred. No. 0.0067;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KEGDKFKLKSELKEL 16

Db 21 REGDKHKLKSELKEL 36

RESULT 12

A48015
S-100 protein beta chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 24-Nov-1999
C:Accession: A48015
R:Jiang, H.; Shah, S.; Hilt, D.C.
J. Biol. Chem. 269, 20502-20511, 1993
A:Title: Organization, sequence, and expression of the murine S100beta gene. Transcription of the murine S100beta gene.
A:Reference number: A48015; MUID:93388628; PMID:8376406
A:Accession: A48015
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-92 <JIA>
A:Cross-references: GB:L22144; NID:g404768; PIDN:AAA03075.1; PID:g404769
C:Genetics:
A:Introns: 46/3
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: acetylated amino end; calcium binding; EF hand
F:2-92/Product: S-100 protein beta chain #status predicted <MAT>
F:6-40/Domain: calmodulin repeat homology <EF1>
F:49-81/Domain: calmodulin repeat homology <EF2>

F:2/Modified site: acetylated amino end (Ser) (in mature form) #status predicted
F:20,22,24,27,32/Binding site: calcium (Gly, Glu, Asp, Lys, Glu) #status predicted
F:62,64,66,68,70,73/Binding site: calcium (Asp, Asp, Glu, Asp, Glu) #status predicted

Query Match 80.0%; Score 64; DB 2; Length 92;

Best Local Similarity 81.2%; Pred. No. 0.0067;

Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KEGDKFKLKSELKEL 16

Db 21 REGDKHKLKSELKEL 36

RESULT 13

S35985
S-100 protein alpha chain - weatherfish
C:Species: Misgurnus fossilis (weatherfish)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S35985
R:Rivanekov, V.V.; Gerke, V.; Minin, A.A.; Plessmann, U.; Weber, K.
Mech. Dev. 42, 151-158, 1993
A:Title: Transduction of Ca(2+) signals upon fertilization of eggs; identification of an alpha subunit of the S-100 protein.
A:Reference number: S35985; MUID:94031845; PMID:8217841
A:Accession: S35985
A:Molecule type: protein
A:Residues: 1-95 <IVA>
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; EF hand
F:49-81/Domain: calmodulin repeat homology <EF2>

Query Match 80.0%; Score 64; DB 1; Length 95;

Best Local Similarity 75.0%; Pred. No. 0.0069;

Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KEGDKFKLKSELKEL 16

Db 21 KEGDKYKLSKSELKSL 36

RESULT 14

A54314
calyculin - mouse
C:Species: Mus musculus (house mouse)
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 13-Aug-1999
C:Accession: A54314; S22084; A54559; A49738; S14090; S06647; S10445
R:Wood, L.; Carter, D.; Mills, M.; Hatzenbuehler, N.; Vogeli, G.
J. Invest. Dermatol. 96, 383-387, 1991
A:Title: Expression of calyculin, a calcium-binding protein, in the keratogenous region (keratinocytes) of the epidermis.
A:Reference number: A54314; MUID:91162002; PMID:2002257
A:Accession: A54314
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-89 <WOO>
A:Experimental source: strain 129
R:Timmons, P.; Chan, C.T.J.; Rigby, P.W.J.; Poirier, F.
submitted to the EMBL Data Library, April 1992
A:Description: The gene encoding the calcium binding protein, calyculin, is expressed at the keratogenous region of the epidermis.
A:Reference number: S22084
A:Accession: S22084
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-89 <TIM>
A:Cross-references: EMBL:X66449; NID:g50271; PIDN:CAA47065.1; PID:g50272
R:Guo, X.; Chambers, A.F.; Parfett, C.L.J.; Waterhouse, P.; Murphy, L.C.; Reid, R.E.; Crawford, J.; Chambers, A.F.; Parfett, C.L.J.; Waterhouse, P.; Murphy, L.C.; Reid, R.E.; Crawford, J.
Cell Growth Differ. 1, 333-338, 1990
A:Title: Identification of a serum-inducible messenger RNA (SEI10) as the mouse homologue of the human SEI10.
A:Reference number: A54559; MUID:91120152; PMID:2177633
A:Accession: A54559
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-50 'R', 52-89 <GUO>
A:Cross-references: EMBL:X52278; NID:g49752; PIDN:CAA36521.1; PID:g49753
R:Thordarson, G.; Southard, J.N.; Talamantes, F.

Thu Sep 9 12:03:23 2004

Endocrinology 129, 1257-1265, 1991
A:Title: Purification and characterization of mouse decidua calycyclin: a novel stimulat
A:Reference number: A49738; MUID:91339739; PMID:1874170
A:Accession: A49738
A>Status: preliminary
A:Molecule type: protein
A:Residues: 24-33;37-40;42-51;53-61,'X',63-89 <THO>
R:Filipek, A.; Gerke, V.; Weber, K.; Kuznicki, J.
Eur. J. Biochem. 195, 795-800, 1991
A:Title: Characterization of the cell-cycle-regulated protein calycyclin from Ehrlich asc
A:Reference number: S14090; MUID:91153321; PMID:1999197
A:Accession: S14090
A:Molecule type: protein
A:Residues: 'XXX',6-34,'X',36-89,'X' <FIL>
R:Kuznicki, J.; Filipek, A.; Hunziker, P.E.; Huber, S.; Heizmann, C.W.
Biochem. J. 263, 951-956, 1989
A:Title: Calcium-binding protein from mouse Ehrlich ascites-tumour cells is homologous t
A:Reference number: S06647; MUID:90088446; PMID:2597136
A:Accession: S06647
A:Molecule type: protein
A:Residues: 56-69 <KUZ>
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; EF hand
F:7-41/Domain: calmodulin repeat homology <EF1>
F:48-80/Domain: calmodulin repeat homology <EF2>

Query Match 72.5%; Score 58; DB 2; Length 89;
Best Local Similarity 75.0%; Pred. No. 0.055;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
DB 22 KEGDKHTLSKKELKEL 37

Query Match 72.5%; Score 58; DB 2; Length 89;
Best Local Similarity 75.0%; Pred. No. 0.055;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

RESULT 15
S27011
calycyclin - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: S27011; S20264; S16116
R:Ando, Y.; Watanabe, M.; Akatsuka, H.; Tokumitsu, H.; Hidaka, H.
FEBS Lett. 314, 109-113, 1992

A:Title: Site-directed mutation makes rabbit calycyclin dimer.
A:Reference number: S27011; MUID:93093152; PMID:1459239
A:Accession: S27011
A:Molecule type: mRNA
A:Residues: 1-90 <AND>
A:Cross-references: EMBL:D10885; NID:g217743; PIDN:BRA01707.1; PID:g217744
A:Experimental source: lung
R:Tokumitsu, H.; Kobayashi, R.; Hidaka, H.
Arch. Biochem. Biophys. 291, 401, 1991

A:Title: Corrigendum. A calcium-binding protein from rabbit lung cytosol identified as b
A:Reference number: S20264; MUID:92061074; PMID:1952954
A:Accession: S20264
A:Molecule type: protein
A:Residues: 23-30;35-46;55-89 <TOK>
A:Experimental source: lung
R:Tokumitsu, H.; Kobayashi, R.; Hidaka, H.
Arch. Biochem. Biophys. 288, 202-207, 1991

A:Title: A calcium-binding protein from rabbit lung cytosol identified as the product of
A:Reference number: S16116; MUID:91378440; PMID:1898017
A:Accession: S16116
A:Molecule type: protein
A:Residues: 23-27,'G',29-30;35-42,'G',44-46;55-81,'V',83-89 <TOW>
A:Experimental source: lung
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; EF hand; monomer
F:7-41/Domain: calmodulin repeat homology <EF1>
F:48-80/Domain: calmodulin repeat homology <EF2>

Query Match 72.5%; Score 58; DB 1; Length 90;
Best Local Similarity 75.0%; Pred. No. 0.055;

Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 1 KEGDKFKLNKSELKEL 16
DB 22 KEGDKHTLSKKELKEL 37

Query Match 72.5%; Score 58; DB 2; Length 89;
Best Local Similarity 75.0%; Pred. No. 0.055;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
DB 22 KEGDKHTLSKKELKEL 37

Query Match 72.5%; Score 58; DB 2; Length 89;
Best Local Similarity 75.0%; Pred. No. 0.055;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Search completed: September 9, 2004, 10:59:39
Job time : 9.15484 secs

Search completed: September 9, 2004, 10:59:39
Job time : 9.15484 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:53:18 ; Search time 21.471 Seconds
(without alignments)
238.975 Million cell updates/sec

Title: US-10-067-618-5
Perfect score: 80
Sequence: 1 KEGDKFKLNKSELKEL 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Length	DB ID	Description
1	80	100.0	16	13	US-10-067-618-5
2	80	100.0	16	13	US-10-135-152-5
3	80	100.0	101	9	US-09-393-433-1
4	80	100.0	101	9	US-09-781-509-1
5	80	100.0	101	12	US-10-087-192-1158
6	80	100.0	101	13	US-10-067-618-2
7	80	100.0	101	13	US-10-135-152-2
8	80	100.0	101	14	US-10-269-643-1
9	77	96.2	101	9	US-09-393-433-2
10	77	96.2	101	9	US-09-781-509-2
11	77	96.2	101	12	US-10-087-192-1155
12	77	96.2	119	12	US-10-087-192-1155
13	68	85.0	94	14	US-10-097-340-270
14	65	81.2	96	14	US-10-171-311-204
15	65	81.2	97	14	US-10-097-340-274

16	65	81.2	97	14	US-10-171-311-206	Sequence 206, App
17	65	81.2	97	15	US-10-236-031B-40	Sequence 40, Appl
18	64	80.0	66	12	US-10-336-603A-104	Sequence 104, App
19	64	80.0	92	10	US-09-492-026-5	Sequence 5, Appl
20	64	80.0	92	10	US-09-919-039-184	Sequence 184, App
21	64	80.0	92	12	US-10-336-603A-102	Sequence 102, App
22	64	80.0	97	16	US-10-363-829-412	Sequence 412, App
23	55	68.8	50	14	US-10-029-386-30749	Sequence 30749, A
24	55	68.8	73	12	US-10-424-599-233903	Sequence 233903
25	55	68.8	90	9	US-09-738-973-200	Sequence 200, App
26	55	68.8	90	9	US-09-974-298-20	Sequence 20, Appl
27	55	68.8	90	9	US-09-854-133-200	Sequence 200, App
28	55	68.8	90	14	US-10-097-340-276	Sequence 276, App
29	55	68.8	90	14	US-10-144-649A-200	Sequence 200, App
30	55	68.8	105	14	US-10-106-698-4570	Sequence 4570, App
31	55	68.8	105	15	US-10-264-049-2219	Sequence 2219, App
32	54	67.5	90	16	US-10-408-765A-2429	Sequence 2429, App
33	53	66.2	44	10	US-09-877-843-84	Sequence 84, Appl
34	47	58.8	603	12	US-09-972-211-102	Sequence 102, App
35	47	58.8	603	12	US-10-036-625-102	Sequence 102, App
36	45	56.2	480	15	US-10-369-493-10444	Sequence 10444, A
37	45	56.2	1225	15	US-10-369-493-21875	Sequence 21875, A
38	43	53.8	398	15	US-10-104-047-3388	Sequence 3388, App
39	43	53.8	511	15	US-10-015-115-46	Sequence 46, Appl
40	43	53.8	1173	15	US-10-369-493-21954	Sequence 21954, A
41	42	52.5	381	15	US-10-369-493-11274	Sequence 11274, A
42	42	52.5	597	12	US-10-335-977-8789	Sequence 8789, App
43	42	52.5	621	12	US-10-335-977-8790	Sequence 8790, App
44	42	52.5	634	12	US-10-335-977-8791	Sequence 8791, App
45	42	52.5	638	12	US-10-335-977-8792	Sequence 8792, App

ALIGNMENTS

RESULT 1
US-10-067-618-5
; Sequence 5, Application US/10067618
; Publication No. US20020169110A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 7879ZTA111-Z
; CURRENT APPLICATION NUMBER: US/10/067,618
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/298,625
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide
US-10-067-618-5

Query Match 100.0%; Score 80; DB 13; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
|||||
Db 1 KEGDKFKLNKSELKEL 16

RESULT 2
US-10-135-152-5
; Sequence 5, Application US/10135152
; Publication No. US20020172680A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/09/781,509
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 09/393,433
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-781-509-1
Query Match 100.0%; Score 80; DB 9; Length 101;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KEGDKFKLNKSELKEL 16
Db 22 KEGDKFKLNKSELKEL 37
RESULT 5
US-10-087-192-1158
; Sequence 1158, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1158
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-1158
Query Match 100.0%; Score 80; DB 12; Length 101;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KEGDKFKLNKSELKEL 16
Db 22 KEGDKFKLNKSELKEL 37
RESULT 6
US-10-067-618-2
; Sequence 2, Application US/10067618
; Publication No. US20020169110A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 78792YAI11-Z
; CURRENT APPLICATION NUMBER: US/10/067,618
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/298,625
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 08/468,942

US-10-135-152-5
; Sequence 5, Application US/10135152
; Publication No. US20020172680A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 78792YAI11-Z
; CURRENT APPLICATION NUMBER: US/10/135,152
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US/09/298,625
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide
US-10-135-152-5
Query Match 100.0%; Score 80; DB 13; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KEGDKFKLNKSELKEL 16
Db 1 KEGDKFKLNKSELKEL 16
RESULT 3
US-09-393-433-1
; Sequence 1, Application US/09393433
; Patent No. US20010011126A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/09/393,433
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-393-433-1
Query Match 100.0%; Score 80; DB 9; Length 101;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KEGDKFKLNKSELKEL 16
Db 22 KEGDKFKLNKSELKEL 37
RESULT 4
US-09-781-509-1
; Sequence 1, Application US/09781509
; Patent No. US20020099010A1

```
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-067-618-2

Query Match      100.0%; Score 80; DB 13; Length 101;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
Db 22 KEGDKFKLNKSELKEL 37

RESULT 7
US-10-135-152-2
; Sequence 2, Application US/10135152
; Publication No. US20020172680A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 7879ZVAII-Z
; CURRENT APPLICATION NUMBER: US/10/135,152
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US/09/298,625
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-135-152-2

Query Match      100.0%; Score 80; DB 13; Length 101;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
Db 22 KEGDKFKLNKSELKEL 37

RESULT 8
US-10-269-643-1
; Sequence 1, Application US/10269643
; Publication No. US20030100503A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/10/269,643
; CURRENT FILING DATE: 2002-10-11

Query Match      100.0%; Score 80; DB 13; Length 101;
Best Local Similarity 100.0%; Pred. No. 7.8e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
Db 22 KEGDKFKLNKSELKEL 37

RESULT 9
US-09-393-433-2
; Sequence 2, Application US/09393433
; Patent No. US20010011126A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/09/393,433
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-393-433-2

Query Match      96.2%; Score 77; DB 9; Length 101;
Best Local Similarity 93.8%; Pred. No. 0.00023;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
Db 22 KEGDKFKLNKTELKEL 37

RESULT 10
US-09-781-509-2
; Sequence 2, Application US/09781509
; Patent No. US20020099010A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/09/781,509
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 09/393,433
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-781-509-2

Query Match      96.2%; Score 77; DB 9; Length 101;
```

Best Local Similarity 93.8%; Pred. No. 0.00023; DB 14; Length 101;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
| | | | | | | | | | | | | | | | | | | | | |
DB 22 KEGDKFKLNKSELKEL 37

RESULT 11
US-10-269-643-2
; Sequence 2, Application US/10269643
; Publication No. US20030100503A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/10/269,643
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: US/09/781,509
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 09/393,433
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-269-643-2

Query Match 96.2%; Score 77; DB 14; Length 101;
Best Local Similarity 93.8%; Pred. No. 0.00023;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
| | | | | | | | | | | | | | | | | | | | | |
DB 22 KEGDKFKLNKSELKEL 37

RESULT 12
US-10-087-192-1155
; Sequence 1155, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-1155

Query Match 96.2%; Score 77; DB 12; Length 119;
Best Local Similarity 93.8%; Pred. No. 0.00027;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
| | | | | | | | | | | | | | | | | | | | | |
DB 40 KEGDKFKLNKSELKEL 55

RESULT 13
US-10-097-340-270
; Sequence 270, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPURU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 270
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-270

Query Match 85.0%; Score 68; DB 14; Length 94;
Best Local Similarity 81.2%; Pred. No. 0.0052;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
| | | | | | | | | | | | | | | | | | | | | |
DB 22 KEGDKFKLNKSELKEL 37

RESULT 14
US-10-171-311-204
; Sequence 204, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersch, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY

```
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; PRIOR FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 204
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-204

Query Match      81.2%; Score 65; DB 14; Length 96;
Best Local Similarity 75.0%; Pred. No. 0.016;
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY      1 KEGDKFKLNKSELKEL 16
       :|||||:|:|:|
Db      23 QEGDKFKLSKGEMKEL 38

RESULT 15
US-10-097-340-274
; Sequence 274, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPURU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
```

US-10-097-340-274

Query Match 81.2%; Score 65; DB 14; Length 97;
Best Local Similarity 75.0%; Pred. No. 0.016;
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
 :|||||:|:|:|
Db 22 QEGDKFKLSKGEMKEL 37

Search completed: September 9, 2004, 11:04:46
Job time : 22.471 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:47:43 : Search time 8.98065 Seconds
(without alignments)
91.977 Million cell updates/sec

Title: US-10-067-618-5

Perfect score: 80

Sequence: 1 KEGDKFKLNKSELKEL 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pap.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pap.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pap.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pap.*
5: /cgn2_6/ptodata/2/iaa/PCTUS COMB.pap.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	16	US-08-190-560-5	Sequence 5, Appli
2	80	100.0	16	US-08-469-277-5	Sequence 5, Appli
3	80	100.0	16	US-08-468-946-5	Sequence 5, Appli
4	80	100.0	16	US-08-468-942-5	Sequence 5, Appli
5	80	100.0	16	US-09-298-625-5	Sequence 5, Appli
6	80	100.0	101	US-08-190-560-2	Sequence 2, Appli
7	80	100.0	101	US-08-469-277-2	Sequence 2, Appli
8	80	100.0	101	US-08-468-946-2	Sequence 2, Appli
9	80	100.0	101	US-08-468-942-2	Sequence 2, Appli
10	80	100.0	101	US-09-298-625-2	Sequence 2, Appli
11	65	81.2	97	US-07-662-198B-2	Sequence 2, Appli
12	64	80.0	45	US-08-056-200-97	Sequence 2, Appli
13	64	80.0	45	US-08-800-644-97	Sequence 97, Appl
14	64	80.0	91	US-07-987-272A-11	Sequence 91, Appl
15	64	80.0	92	US-08-918-727-5	Sequence 11, Appl
16	64	80.0	92	US-09-051-589-1	Sequence 5, Appli
17	64	80.0	92	US-09-205-680A-5	Sequence 1, Appli
18	55	68.8	44	US-08-056-200-100	Sequence 5, Appli
19	55	68.8	44	US-08-800-644-100	Sequence 100, App
20	55	68.8	88	US-09-621-976-7524	Sequence 100, App
21	55	68.8	88	US-09-621-976-7526	Sequence 7524, Ap
22	55	68.8	88	US-09-621-976-7528	Sequence 7526, Ap
23	55	68.8	89	US-07-987-272A-10	Sequence 7528, Ap
24	55	68.8	90	US-09-370-838-200	Sequence 10, Appl
25	55	68.8	92	US-09-621-976-7537	Sequence 200, App
26	45	56.2	64	US-09-621-976-4767	Sequence 7537, Ap
27	44	55.0	75	US-07-987-272A-12	Sequence 4767, Ap
					Sequence 12, Appl

28 41 51.2 764 1 US-08-375-300-4 Sequence 4, Appli
29 41 51.2 764 3 US-09-177-431-4 Sequence 4, Appli
30 41 51.2 764 5 PCT-US95-16930-4 Sequence 4, Appli
31 41 51.2 1089 1 US-08-375-300-2 Sequence 2, Appli
32 41 51.2 1089 3 US-09-177-431-2 Sequence 2, Appli
33 41 51.2 1089 5 PCT-US95-16930-2 Sequence 2, Appli
34 40 50.0 181 4 US-09-134-000C-3749 Sequence 3749, Ap
35 40 50.0 711 3 US-08-949-588-2 Sequence 2, Appli
36 39.5 49.4 205 4 US-08-198-452A-829 Sequence 829, App
37 39 48.8 320 4 US-08-956-171E-5192 Sequence 5192, Ap
38 39 48.8 604 4 US-09-519-232-64 Sequence 64, Appl
39 39 48.8 755 4 US-09-360-545-16 Sequence 6691, Ap
40 39 48.8 782 4 US-09-398-395A-46 Sequence 16, Appl
41 39 48.8 782 4 US-09-887-586A-46 Sequence 46, Appl
42 39 48.8 782 4 US-09-895-752-46 Sequence 46, Appl
43 39 48.8 782 4 US-09-903-012B-46 Sequence 46, Appl
44 39 48.8 782 4 US-09-900-797-46 Sequence 46, Appl
45 39 48.8 782 4 US-09-900-797-46 Sequence 46, Appl

ALIGNMENTS

RESULT 1
US-08-190-560-5
; Sequence 5, Application US/08190560
; Patent No. 5798257
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-190-560-5

Query Match 100.0%; Score 80; DB 1; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
Db 1 KEGDKFKLNKSELKEL 16

```

1  ; ZIP: 11530
2  ; COMPUTER READABLE FORM:
3  ; MEDIUM TYPE: Floppy disk
4  ; COMPUTER: IBM PC compatible
5  ; OPERATING SYSTEM: PC-DOS/MS-DOS
6  ; SOFTWARE: PatentIn Release #1.0, Version #1.25
7  ; CURRENT APPLICATION DATA:
8  ; APPLICATION NUMBER: US/08/468,946
9  ; FILING DATE: 06-JUN-1995
10 ; CLASSIFICATION: 435
11 ; PRIOR APPLICATION DATA:
12 ; APPLICATION NUMBER: US 08/190,560
13 ; FILING DATE: 31-JAN-1994
14 ; ATTORNEY/AGENT INFORMATION:
15 ; NAME: Digiglio, Frank S.
16 ; REGISTRATION NUMBER: 31,346
17 ; REFERENCE/DOCKET NUMBER: 78792Y
18 ; TELECOMMUNICATION INFORMATION:
19 ; TELEPHONE: (516) 742-4343
20 ; TELEFAX: (516) 742-4366
21 ; TELEX: 230 901 SANS UR
22 ; INFORMATION FOR SEQ ID NO: 5:
23 ; SEQUENCE CHARACTERISTICS:
24 ; LENGTH: 16 amino acids
25 ; TYPE: amino acid
26 ; STRANDEDNESS: double
27 ; TOPOLOGY: linear
28 ; MOLECULE TYPE: DNA (genomic)
29 ; US-08-468-946-5
30 ;
31 ; Query Match 100.0%; Score 80; DB 2; Length 16;
32 ; Best Local Similarity 100.0%; Pred. No. 1.7e-06;
33 ; Matches 16; Conservative 0; Mismatches 0; Indels
34 ;
35 ; QY 1 KEGDKFKLNKSELKEL 16
36 ; DDB 1 KEGDKFKLNKSELKEL 16
37 ;
38 ; RESULT 4
39 ; US-08-468-942-5
40 ; Sequence 5, Application US/08468942
41 ; Patent No. 5965360
42 ; GENERAL INFORMATION:
43 ; APPLICANT: Zain, Sayeeda
44 ; APPLICANT: Lukanidin, Eugene
45 ; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
46 ; TITLE OF INVENTION: THE MTS-1 GENE
47 ; NUMBER OF SEQUENCES: 9
48 ; CORRESPONDENCE ADDRESS:
49 ; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
50 ; STREET: 400 Garden City Plaza
51 ; CITY: Garden City
52 ; STATE: New York
53 ; COUNTRY: United States
54 ; ZIP: 11530
55 ; COMPUTER READABLE FORM:
56 ; MEDIUM TYPE: Floppy disk
57 ; COMPUTER: IBM PC compatible
58 ; OPERATING SYSTEM: PC-DOS/MS-DOS
59 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
60 ; CURRENT APPLICATION DATA:
61 ; APPLICATION NUMBER: US/08/468,942
62 ; FILING DATE: 06-JUN-1995
63 ; CLASSIFICATION: 435
64 ; PRIOR APPLICATION DATA:
65 ; APPLICATION NUMBER: US 08/190,560
66 ; FILING DATE: 31-JAN-1994
67 ; ATTORNEY/AGENT INFORMATION:
68 ; NAME: Digiglio, Frank S.
69 ; REGISTRATION NUMBER: 31,346
70 ; REFERENCE/DOCKET NUMBER: 78792Y
71 ; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-468-942-5

Query Match
Best Local Similarity 100.0%; Score 80; DB 2; Length 16;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
DB 1 KEGDKFKLNKSELKEL 16

RESULT 5
US-09-298-625-5
; Sequence 5, Application US/09298625
; Patent No. 6638504
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 78792YAI11-Z
; CURRENT APPLICATION NUMBER: US/09/298,625
; CURRENT FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide
US-09-298-625-5

Query Match
Best Local Similarity 100.0%; Score 80; DB 4; Length 16;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
DB 1 KEGDKFKLNKSELKEL 16

RESULT 6
US-08-190-560-2
; Sequence 2, Application US/08190560
; Patent No. 5798257
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York

; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-468-942-5

Query Match
Best Local Similarity 100.0%; Score 80; DB 1; Length 101;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
DB 22 KEGDKFKLNKSELKEL 37

RESULT 7
US-08-469-277-2
; Sequence 2, Application US/08469277
; Patent No. 5801142
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,277
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-190-560-2

```

TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-277-2

Query Match 100.0%; Score 80; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
|||
Db 22 KEGDKFKLNKSELKEL 37

RESULT 8
US-08-468-946-2
; Sequence 2, Application US/08468946
; Patent No. 5843696
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,946
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,560
FILING DATE: 31-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 78792Y
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 101 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-946-2

Query Match 100.0%; Score 80; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
|||
Db 22 KEGDKFKLNKSELKEL 37

RESULT 9
US-08-468-942-2
; Sequence 2, Application US/08468942
; Patent No. 5965360
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,942
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,560
FILING DATE: 31-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 78792Y
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 101 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-468-942-2

Query Match 100.0%; Score 80; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
|||
Db 22 KEGDKFKLNKSELKEL 37

RESULT 10
US-09-298-625-2
; Sequence 2, Application US/09298625
; Patent No. 6638504
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 78792YXIII-2
; CURRENT APPLICATION NUMBER: US/09/298,625
; CURRENT FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600

US-09-298-625-2
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens

Query Match 100.0%; Score 80; DB 4; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.1e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KEGDKFKLNKSELKEL 16
Db 22 KEGDKFKLNKSELKEL 37

RESULT 11
US-07-662-198B-2
; Sequence 2, Application US/07662198B
; Patent No. 5262528
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; APPLICANT: Lee, Sam W.
; APPLICANT: Tomasetto, Catherine
; TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: Wordperfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/662.198B
; FILING DATE: 19910228
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/049001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: linear

Query Match 81.2%; Score 65; DB 1; Length 97;
Best Local Similarity 75.0%; Pred. No. 0.0025;
Matches 12; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KEGDKFKLNKSELKEL 16
Db 22 QEGDKFKLNKSELKEL 37

RESULT 12

US-08-056-200-97
; Sequence 97, Application US/08056200
; Patent No. 5616500
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: U.S.A.
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/056.200
; FILING DATE: 30-APR-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fedrick, Michael F.
; REGISTRATION NUMBER: 36,799
; REFERENCE/DOCKET NUMBER: NIH054.001A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (714) 760-0404
; TELEFAX: (714) 760-9502
; INFORMATION FOR SEQ ID NO: 97:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 45 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal

US-08-056-200-97
Query Match 80.0%; Score 64; DB 1; Length 45;
Best Local Similarity 81.2%; Pred. No. 0.0017;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KEGDKFKLNKSELKEL 16
Db 20 REGDKFKLNKSELKEL 35

RESULT 13
US-08-800-644-97
; Sequence 97, Application US/08800644
; Patent No. 5958752
; GENERAL INFORMATION:
; APPLICANT: Steinert, Peter M.
; APPLICANT: Lee, Seung-Chul
; APPLICANT: Kim, In-Gyu
; APPLICANT: Chung, Soo-Il
; APPLICANT: Park, Sang-Chul
; TITLE OF INVENTION: Trichohyalin and Transglutaminase-3 and
; TITLE OF INVENTION: Methods of Using Same
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 620 Newport Center Drive, Sixteenth Floor
; CITY: Newport Beach

STATE: CA
COUNTRY: U.S.A.
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,644
FILING DATE: 14-FEB-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/056,200
FILING DATE: 30-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Redrick, Michael F.
REGISTRATION NUMBER: 36,799
REFERENCE/DOCKET NUMBER: NIH054.001A
TELEPHONE: (714) 760-0404
TELEFAX: (714) 760-9502
INFORMATION FOR SEQ ID NO: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 45 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
US-08-800-644-97

Query Match 80.0%; Score 64; DB 2; Length 45;
Best Local Similarity 81.2%; Pred. No. 0.0017;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
:||||| ||| |||||
Db 20 REGDKHKLKSELKEL 35

RESULT 14
US-07-987-272A-11
Sequence 11, Application US/07987272A
Patent No. 5731166
GENERAL INFORMATION:
APPLICANT: Geczy, C., Simpson, R. J. and Lackmann, M
TITLE OF INVENTION: No. 5731166el Chemotactic factor
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cushman Darby & Cushman
STREET: 1100 New York Avenue, N. W., Ninth Floor, East Tower
CITY: Washington
STATE: D. C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/987,272A
FILING DATE: 05-MAR-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PK 2127
FILING DATE: 05-FEB-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PK 4463
FILING DATE: 05-SEP-1991

ATTORNEY/AGENT INFORMATION:
NAME: Brinkman, David W.
REGISTRATION NUMBER: 20,817
REFERENCE/DOCKET NUMBER: DWB/1925/200259
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861 3000
TELEFAX: 202-822 0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-987-272A-11

Query Match 80.0%; Score 64; DB 1; Length 91;
Best Local Similarity 81.2%; Pred. No. 0.0034;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
:||||| ||| |||||
Db 20 REGDKHKLKSELKEL 35

RESULT 15
US-08-918-727-5
Sequence 5, Application US/08918727
Patent No. 5849528
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Bandman, Olga
APPLICANT: Corley, Neil C.
APPLICANT: Lal, Preeti
APPLICANT: Shah, Purvi
TITLE OF INVENTION: HUMAN S100 PROTEINS
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/918,727
FILING DATE: Herewith
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0373 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 92 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank

CLONE: 337730
US-08-918-727-5

Query Match 80.0%; Score 64; DB 2; Length 92;
Best Local Similarity 81.2%; Pred. No. 0.0035;
Matches 13; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
Db 21 REGDKHKLKSELKEL 36

Search completed: September 9, 2004, 11:01:11
Job time : 9.98065 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:32:01 ; Search time 33.6516 Seconds
(without alignments)
134.340 Million cell updates/sec

Title: US-10-067-618-5

Perfect score: 80

Sequence: 1 KEGDKFKLNKSELKEL 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	80	100.0	16	2 AAR80455	Aar80455 Human mts
2	80	100.0	16	3 AAB37434	Aab37434 Human mts
3	80	100.0	16	6 ABU08519	Abu08519 Human mts
4	80	100.0	101	2 AAR20560	Aar20560 Human mts
5	80	100.0	101	2 AAR80453	Aar80453 Human mts
6	80	100.0	101	3 AAB45534	Aab45534 Human S10
7	80	100.0	101	3 AAB37432	Aab37432 Human mts
8	80	100.0	101	4 AAB72386	Aab72386 Human mts
9	80	100.0	101	6 AAB08513	Abu08513 Human mts
10	80	100.0	101	7 ADD14157	Add14157 Human mts
11	77	96.2	101	4 AAB72387	Aab72387 Murine src
12	68	85.0	93	7 ADE61764	Ade61764 Human S10
13	68	85.0	94	3 AAB45531	Aab45531 Human S10
14	68	85.0	94	4 AAM40258	Aam40258 Human pol
15	68	85.0	94	5 ABG96406	Abg96406 Human ova
16	68	85.0	166	4 AAM42044	Aam42044 Human pol
17	65	81.2	93	7 ADE61762	Ade61762 Rat Prote
18	65	81.2	96	6 ABUS6413	Abu56413 Lung canc
19	65	81.2	96	7 ABR92147	Abu56413 Human cer
20	65	81.2	96	7 AAO30415	Aao30415 Human sec
21	65	81.2	97	2 AAR27058	Aar27058 Sequence
22	65	81.2	97	2 AAR26406	Aar26406 Sequence
23	65	81.2	97	3 AAB45532	Aab45532 Human S10
24	65	81.2	97	5 ABG96408	Abg96408 Human ova
25	65	81.2	97	6 ABR92148	Abu92148 Human cer

ALIGNMENTS

RESULT 1

AAR80455
ID AAR80455 standard; peptide; 16 AA.

XX AC AAR80455;

XX AC AAR80455;

DT 27-DEC-1995 (first entry)

XX Human mts-1 calcium binding domain.

XX Metastasis; cancer; mts-1 gene; tumour; therapy; antigen; antibody.

XX Synthetic.

XX WO9520656-A1.

XX 03-AUG-1995.

XX 31-JAN-1995; 95WO-US001214.

XX 31-JAN-1994; 94US-00190560.

XX (RESE) RESEARCH CORP TECHNOLOGIES INC.

XX Zain S, Lukanidin E;

XX WPI; 1995-275441/36.

XX Nucleic acid encoding human mts-1, antigenic fragments and antibodies -

XX useful for diagnosis of malignant cancer and metastatic potential of

XX tumour cells.

XX Disclosure; Page 24; 124pp; English.

XX Antigenic peptides 1-4 (given in AAR80454-57) of human mts-1 protein

XX (AAR80453) were used to raise polyclonal and monoclonal antibodies that

XX detect the presence of mts-1 in tissue samples, esp. metastatic cells.

XX Peptides 1, 3 and 4 are unique to mts-1, while peptide 2 generates

XX antibodies reactive with many calcium binding proteins

XX Sequence 16 AA;

XX Query Match

XX Best Local Similarity 100.0%; Score 80; DB 2; Length 16;

XX Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 1 KEGDKFKLNKSELKEL 16

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

|||||

Db 1 KEGDKFKLNKSELKEL 16

RESULT 2

AAB37434

ID AAB37434 standard; peptide; 16 AA.

XX AC AAB37434;

XX DT 21-FEB-2001 (first entry)

XX DE Human mts-1 peptide fragment #2.

XX KW Human; mts-1; cytostatic; cancer; metastasis.

XX OS Homo sapiens.

XX PN WO200064475-A1.

XX PD 02-NOV-2000.

XX PF 20-APR-2000; 2000WO-US011006.

XX PR 23-APR-1999; 99US-00298625.

XX PA (RESE) RESEARCH CORP TECHNOLOGIES INC.

XX PI Lukanidin E;

XX DR WPI; 2000-687266/67.

XX PT Treating cancer and preventing metastasis comprises administration of an anti-mts-1 protein antibody or antisense oligonucleotide.

XX PS Example 9; Page 27; 155pp; English.

XX CC The present invention relates to methods for treating cancer and preventing metastases, comprising the administration of a composition directed against the mts-1 protein (see AAC68131-C68132 and AAB37432).

XX CC Mts-1 protein is a calcium-binding protein, and is thought to have a role in myoepithelial cell differentiation. The present sequence is a peptide fragment of the human mts-1 protein. This peptide was used to generate antibodies against mts-1 protein, which can be used to detect mts-1 protein in clinical specimens

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 80; DB 3; Length 16;

Best Local Similarity 100.0%; Pred. No. 9.8e-07; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0;

QY 1 KEGDKFKLNKSELKEL 16

Db 1 KEGDKFKLNKSELKEL 16

RESULT 3

ABU08519

ID ABU08519 standard; peptide; 16 AA.

XX AC ABU08519;

XX DT 27-MAY-2003 (first entry)

XX DE Human mts-1 protein, antigenic peptide #2.

XX KW Human; mts-1; metastatic cancer; p53; cytostatic; gene therapy; antigen.

XX OS Homo sapiens.

XX PN US2002172680-A1.

XX PD 21-NOV-2002.

XX

XX PF 29-APR-2002; 2002US-00135152.

XX PR 09-JUL-1990; 90US-00550600.

XX PR 25-NOV-1992; 92US-00981455.

XX PR 31-JAN-1994; 94US-00190560.

XX PR 06-JUN-1995; 95US-00468942.

XX PR 23-APR-1999; 99US-00298635.

XX PA (LUKA/) LUKANIDIN E.

XX PI Lukanidin E;

XX WPI; 2003-328422/31.

XX PT Treating cancer by administering a reagent directed against the mts-1 protein or an oligonucleotide capable of binding to mts-1 mRNA.

XX PS Example 9; Page 19; 64pp; English.

XX CC The invention relates to treating cancer comprising administering a reagent directed against the mts-1 (not defined) protein or an oligonucleotide capable of binding to mts-1 mRNA. Also include are a method of inactivating, destroying or nullifying a mts-1 protein or cells and a method of inhibiting metastasis in a cancerous cell. Experiments showed that mts-1 inhibited the phosphorylation of full-size p53 and the C-terminal protein fragment by PKC. Addition of the same concentrations of mts-1 to the PKC reaction mixture did not affect the phosphorylation of the N-terminal and DNA-binding domains of p53. The method is useful for treating cancers associated with the mts-1 gene. The present sequence represents a human mts-1 protein antigenic peptide

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 80; DB 6; Length 16;

Best Local Similarity 100.0%; Pred. No. 9.8e-07; Mismatches 0; Indels 0; Gaps 0;

Matches 16; Conservative 0;

QY 1 KEGDKFKLNKSELKEL 16

Db 1 KEGDKFKLNKSELKEL 16

RESULT 4

AAR20560

ID AAR20560 standard; protein; 101 AA.

XX AC AAR20560;

XX DT 25-MAR-2003 (revised)

XX DT 14-MAY-1992 (first entry)

XX DE Human mts protein.

XX KW Metastatic cancer; antibodies; mouse; lung; liver; kidney; thyroid; breast cancer; cell growth.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Protein 1..101 /label= mts-1

XX FT Peptide 2..11 /note= "antigenic"

XX FT Peptide 22..37 /note= "antigenic; calcium binding domain"

XX FT Peptide 42..54 /note= "antigenic"

XX FT Peptide 87..101 /note= "antigenic"

XX PN WO9200757-A.

XX

PD 23-JAN-1992.
 XX
 PF 09-JUL-1990; 90US-00550600.
 XX
 PR 09-JUL-1990; 90US-00550600.
 XX
 PA (RESE) RESEARCH CORP TECHNOLOGIES INC.
 XX
 PI Zain S, Lukanidin E;
 XX
 DR WPI; 1992-056647/07.
 XX
 DR N-PSDB; AAQ20506.
 XX
 XX Metastatic cancer diagnosis by detection of mts-1 gene or protein - using
 PT antibody treatment of cancer and tumours of e.g. kidney, thyroid, lung
 PT and liver.
 XX
 PS Claim 9; Fig 2; 82pp; English.
 XX
 CC The sequence was deduced from the DNA sequence obtd. by screening a human
 CC cDNA library with mouse mts-1 cDNA probes. The antigenic Mts-1 peptides
 CC (see features) derived from the protein and anti- bodies raised to them
 CC are useful in the diagnosis of metastatic cancer, e.g. lung, kidney,
 CC thyroid or breast cancer. The peptide comprising the calcium binding
 CC protein family; the other three peptides are unique to mts-1 and generate
 CC antibodies specific only for this protein. Cell lines capable of
 CC expressing mts-1 are useful as model systems for in vitro and in vivo
 CC anti-metastasis drug screening. Pharmaceutical compns. contg. the mts-1
 CC protein or anti-cancer reagents may be used to promote cell growth, or
 CC for treating cancer, respectively. (Updated on 25-MAR-2003 to correct PA
 CC field.)
 XX
 XX Sequence 101 AA;
 SQ
 Query Match 100.0%; Score 80; DB 2; Length 101;
 Best Local Similarity 100.0%; Pred. No. 8.4e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KEGDKFKLNKSELKEL 16
 DB 22 KEGDKFKLNKSELKEL 37
 RESULT 5
 AAR80453
 ID AAR80453 standard; protein; 101 AA.
 AC AAR80453;
 XX
 DT 27-DEC-1995 (first entry)
 DE Human mts-1 protein.
 XX
 KW Metastasis; cancer; mts-1 gene; tumour; therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9520656-A1.
 XX
 PD 03-AUG-1995.
 XX
 PF 31-JAN-1995; 95WO-US001214.
 XX
 PR 31-JAN-1994; 94US-00190560.
 XX
 PA (RESE) RESEARCH CORP TECHNOLOGIES INC.
 XX
 PI Zain S, Lukanidin E;
 XX
 DR WPI; 1995-275441/36.
 XX
 DR N-PSDB; AAQ99177.
 XX

PT Nucleic acid encoding human mts-1, antigenic fragments and antibodies -
 PT useful for diagnosis of malignant cancer and metastatic potential of
 PT tumour cells.
 XX
 PS Claim 12; Page 92; 124pp; English.
 XX
 CC A human cDNA library was constructed in lambda-gt10 using poly(A) + RNA
 CC prep'd. from HeLa cells. The library was screened with a 32P- labeled
 CC mouse mts-1 vDNA probe. A clone was obtd. which comprised the full-length
 CC human mts-1 gene. The encoded protein is used for the diagnosis or
 CC therapy of cancer, and to raise antibodies
 XX
 SQ Sequence 101 AA;
 Query Match 100.0%; Score 80; DB 2; Length 101;
 Best Local Similarity 100.0%; Pred. No. 8.4e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KEGDKFKLNKSELKEL 16
 DB 22 KEGDKFKLNKSELKEL 37
 RESULT 6
 AAB45534
 ID AAB45534 standard; protein; 101 AA.
 XX
 AC AAB45534;
 XX
 DT 22-FEB-2001 (first entry)
 XX
 DE Human S100A4 protein.
 XX
 KW S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;
 KW calcium-binding protein; calcium homeostasis; cardiac muscle;
 KW pumping capacity; myocardial cell; systolic calcium ion release;
 KW sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;
 KW valve defect.
 XX
 OS Homo sapiens.
 XX
 PN DE19915485-A1.
 XX
 PD 19-OCT-2000.
 XX
 PF 07-APR-1999; 99DE-01015485.
 XX
 PR 07-APR-1999; 99DE-01015485.
 XX
 PA (KATU/) KATUS H A.
 PA (REMP/) REMPPIS A.
 XX
 PI Katus HA, Remppis A;
 XX
 DR WPI; 2000-673510/66.
 XX
 DR N-PSDB; AAC81804.
 XX
 CC Composition containing S100 protein, corresponding nucleic acid or
 CC vector, useful for treating cardiomyopathy and cardiac insufficiency.
 PS
 PS Claim 35; Page 11; 36pp; German.
 XX
 CC This invention describes a novel composition for treating primary or
 CC secondary cardiomyopathy or cardiac insufficiency contains at least one
 CC S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or
 CC fragments, or a gene transfer vector containing (II), optionally
 CC formulated with auxiliaries and/or carriers. (I) are calcium-binding
 CC proteins involved in calcium homeostasis, so their overexpression in
 CC cardiac muscle will improve pumping capacity (and overall capacity) of
 CC the heart. In cultured myocardial cells they increase the contraction and
 CC relaxation rates associated with increased systolic calcium ion release
 CC from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are
 CC used to treat cardiomyopathy (CMP) where inherited or caused by

Thu Sep 9 12:03:22 2004

us-10-067-618-5.rag

CC spontaneous mutations and ischemic CMP caused by arteriosclerosis,
 CC dilative CMP caused by toxic/infectious disease, cardiac disease caused
 CC by pulmonary and/or arterial hypertension, and structural disease caused
 CC by rhythm disorders or valve defects, generally any condition associated
 CC with reduced contractile force. Unlike calmodulin, which is expressed
 CC ubiquitously, (1) show tissue-specific expression and treat the
 CC underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac
 CC disease
 CC XX
 SQ Sequence 101 AA;
 Query Match 100.0%; Score 80; DB 3; Length 101;
 Best Local Similarity 100.0%; Pred. No. 8.4e-06; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 0;
 QY 1 KEGDKFKLNKSELKEL 16
 Db 22 KEGDKFKLNKSELKEL 37
 RESULT 7
 AAB37432
 ID AAB37432 standard; protein; 101 AA.
 XX AC AAB37432;
 XX DT 21-FEB-2001 (first entry)
 XX DE Human mts-1.
 XX KW Human; mts-1; cytostatic; cancer; metastasis.
 XX OS Homo sapiens.
 XX PN WO200064475-A1.
 XX PD 02-NOV-2000.
 XX PF 20-APR-2000; 2000WO-US011006.
 XX PR 23-APR-1999; 99US-00298625.
 XX PA (RESE) RESEARCH CORP TECHNOLOGIES INC.
 XX PI Lukanidin E;
 XX PS WPI; 2000-687266/67.
 XX DR N-PSDB; AAC68131, AAC68132.
 XX PT Treating cancer and preventing metastasis comprises administration of an
 PT anti-mts-1 protein antibody or antisense oligonucleotide.
 XX Example 6; Fig 2; 155pp; English.
 PS The present sequence is human mts-1. Mts-1 protein is a calcium-binding
 CC protein, and is thought to have a role in myoepithelial cell
 CC differentiation. The present invention relates to methods for treating
 CC cancer and preventing metastases, comprising the administration of a
 CC composition directed against the mts-1 protein
 CC XX
 SQ Sequence 101 AA;
 Query Match 100.0%; Score 80; DB 3; Length 101;
 Best Local Similarity 100.0%; Pred. No. 8.4e-06; Indels 0; Gaps 0;
 Matches 16; Conservative 0; Mismatches 0;
 QY 1 KEGDKFKLNKSELKEL 16
 Db 22 KEGDKFKLNKSELKEL 37
 RESULT 8
 AAB72386

ID AAB72386 standard; protein; 101 AA.
 XX AC AAB72386;
 XX DT 24-MAY-2001 (first entry)
 XX DE Human Mts 1 protein amino acid sequence.
 XX KW Mts 1; Mts1-del75; Mts1-4S; S100A4; neuronal cell growth; cardiac arrest;
 KW neuronal degeneration; Parkinson's disease; Alzheimer's disease;
 KW Down's syndrome; stroke; sciatic crush; spinal cord injury; human.
 XX OS Homo sapiens.
 XX PN WO200118043-A2.
 XX PD 15-MAR-2001.
 XX PF 07-SEP-2000; 2000WO-US024495.
 XX PR 10-SEP-1999; 99US-00393433.
 XX PA (PROL-) PROLIFIA INC.
 XX PI Bock E, Lukanidin EM, Berezin V;
 XX PS WPI; 2001-235188/24.
 XX PT New isolated functional derivatives of Mts protein for stimulating
 PT neuronal cell growth to treat neurological conditions e.g. Parkinson's
 PT disease, Alzheimer's disease, Down's syndrome and stroke.
 XX DR Disclosure; Page; 60pp; English.
 XX CC This invention relates to functional derivatives of an Mts protein,
 CC particularly Mts1-del75, and Mts1-4S. Mts 1 (also known as S100A4) is a
 CC member of the S100 gene family. Mts1 proteins and their functional
 CC derivatives are used to stimulate the growth of neuronal cells and in the
 CC treatment of neurological conditions including those characterised by
 CC neuronal degeneration, death or injury such as Parkinson's disease,
 CC Alzheimer's disease, Down's syndrome, stroke, cardiac arrest, sciatic
 CC crush, spinal cord injury, injury to sensory neurons or degenerative
 CC disease of the retina. The present sequence represents the human Mts 1
 CC protein, from which the Mts proteins of the invention may be derived.
 CC NOTE: The present sequence is not shown in the specification but it has
 CC been derived from the sequence given in US patent 5801142 (referred to on
 CC page 10 of the specification)
 XX SQ Sequence 101 AA;
 Query Match 100.0%; Score 80; DB 4; Length 101;
 Best Local Similarity 100.0%; Pred. No. 8.4e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 KEGDKFKLNKSELKEL 16
 Db 22 KEGDKFKLNKSELKEL 37
 RESULT 9
 ABU08513
 ID ABU08513 standard; protein; 101 AA.
 XX AC ABU08513;
 XX DT 27-MAY-2003 (first entry)
 XX DE Human mts-1 protein.
 XX KW Human; mts-1; metastatic cancer; p53; cytostatic; gene therapy.
 XX OS Homo sapiens.
 XX

PN US2002172680-A1.
 PD 21-NOV-2002.
 PF 29-APR-2002; 2002US-00135152.
 PP 09-JUL-1990; 90US-00550600.
 PR 25-NOV-1992; 92US-00981455.
 PR 31-JAN-1994; 94US-00190560.
 PR 06-JUN-1995; 95US-00468942.
 PR 23-APR-1999; 99US-00298635.
 PP (LUKA/) LUKANIDIN E.
 PA Lukanidin E;
 XX WPI; 2003-328422/31.
 DR N-PSDB; ABX93578.
 XX
 PT Treating cancer by administering a reagent directed against the mts-1
 PT protein or an oligonucleotide capable of binding to mts-1 mRNA.
 XX
 PS Example 6; Fig 2; 64pp; English.
 XX
 CC The invention relates to treating cancer comprising administering a
 CC reagent directed against the mts-1 (not defined) protein or an
 CC oligonucleotide capable of binding to mts-1 mRNA. Also include are a
 CC method of inactivating, destroying or nullifying a mts-1 protein or cells
 CC and a method of inhibiting metastasis in a cancerous cell. Experiments
 CC showed that mts-1 inhibited the phosphorylation of full-size p53 and the
 CC C-terminal protein fragment by PKC. Addition of the same concentrations
 CC of mts-1 to the PKC reaction mixture did not affect the phosphorylation
 CC of the N-terminal and DNA-binding domains of p53. The method is useful
 CC for treating cancers associated with the mts-1 gene. The present sequence
 CC represents human mts-1 protein
 XX
 SQ Sequence 101 AA;
 Query Match 100.0%; Score 80; DB 6; Length 101;
 Best Local Similarity 100.0%; Pred. No. 8.4e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KEGDKFKLNKSELKEL 16
 Db 22 KEGDKFKLNKSELKEL 37
 RESULT 10
 ADD14157
 ID ADD14157 standard; protein; 101 AA.
 XX
 AC ADD14157;
 XX
 DT 01-JAN-2004 (first entry)
 DE Human src biomarker polypeptide SEQ ID NO:346.
 XX
 KW predictor set; protein tyrosine kinase activity modulator;
 KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
 KW gene therapy; drug sensitivity; genetic profile; cancer; human.
 XX
 OS Homo sapiens.
 XX
 KW WO2003062395-A2.
 PN
 PP 31-JUL-2003.
 PD
 XX 17-JAN-2003; 2003WO-US001981.
 PF
 XX 18-JAN-2002; 2002US-0350061P.
 PR
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 PA

PI Huang F, Fairchild CR, Lee FY, Shaw P;
 XX WPI; 2003-636735/60.
 DR N-PSDB; ADD14760.
 XX
 PT New polynucleotides and polypeptides for predicting the activity of
 PT compounds that interact with protein tyrosine kinases and/or protein
 PT tyrosine kinase pathways.
 XX
 XX Claim 10; SEQ ID NO 346; 139pp; English.
 PS
 CC The present invention describes a predictor set comprising a plurality of
 CC polynucleotides or polypeptides whose expression pattern is predictive of
 CC the response of cells to treatment with a compound that modulates protein
 CC tyrosine kinase activity or members of the protein tyrosine kinase
 CC pathway. Also described: (1) predicting whether a compound is capable of
 CC modulating the activity of cells, comprising obtaining a sample of cells,
 CC determining whether the cells express a plurality of markers, and
 CC correlating the expression of the markers to the compound's ability to
 CC modulate the activity of the cells; (2) a plurality of cell lines for
 CC identifying polynucleotides and polypeptides whose expression levels
 CC correlate with compound sensitivity or resistance of cells associated
 CC with a disease state; and (3) identifying polynucleotides and
 CC polypeptides that predict compound sensitivity or resistance of cells
 CC associated with a disease state, comprising subjecting the plurality of
 CC cell lines to one or more compounds, analysing the expression pattern of
 CC a microarray of polynucleotides or polypeptides, and selecting
 CC polynucleotides or polypeptides that predict the sensitivity or
 CC resistance of cells associated with a disease state by using the
 CC expression pattern of the microarray. The polynucleotides and
 CC polypeptides have cytostatic activities, and can be used in gene therapy.
 CC The polynucleotides and polypeptides are useful in predicting the
 CC activity of compounds that interact with protein tyrosine kinases and/or
 CC protein tyrosine kinase pathways. These may be used in determining drug
 CC sensitivity in patients to allow the development of individualized
 CC genetic profiles which aid in treating diseases and disorders (e.g.
 CC cancer) based on patient response at a molecular level. The present
 CC sequence is used in the exemplification of the present invention.
 XX
 SQ Sequence 101 AA;
 Query Match 100.0%; Score 80; DB 7; Length 101;
 Best Local Similarity 100.0%; Pred. No. 8.4e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KEGDKFKLNKSELKEL 16
 Db 22 KEGDKFKLNKSELKEL 37
 RESULT 11
 AAB72387
 ID AAB72387 standard; protein; 101 AA.
 XX
 AC AAB72387;
 XX
 DT 24-MAY-2001 (first entry)
 DE Murine Mts 1 protein amino acid sequence.
 XX
 KW Mts 1; Mts1-del75; Mts1-4S; S100A4; neuronal cell growth; cardiac arrest;
 KW neuronal degeneration; Parkinson's disease; Alzheimer's disease;
 KW Down's syndrome; stroke; sciatic crush; spinal cord injury; mouse.
 XX
 OS Mus sp.
 XX
 KW WO200118043-A2.
 PN
 PP 15-MAR-2001.
 PD
 XX 07-SEP-2000; 2000WO-US024495.
 PF
 XX 10-SEP-1999; 99US-00393433.
 PR

XX (PROL-) PROLIFIA INC.
XX Bock E, Lukanidin EM, Berezin V;
XX WPI; 2001-235188/24.
XX New isolated functional derivatives of Mts protein for stimulating
PT neuronal cell growth to treat neurological conditions e.g. Parkinson's
PT disease, Alzheimer's disease, Down's syndrome and stroke.
XX
XX Disclosure; Page; 60pp; English.
XX
XX This invention relates to functional derivatives of an Mts protein,
CC particularly Mts1-del175, and Mts1-48. Mts1 (also known as S100A4) is a
CC member of the S100 gene family. Mts1 proteins and their functional
CC derivatives are used to stimulate the growth of neuronal cells and in the
CC treatment of neurological conditions including those characterised by
CC neuronal degeneration, death or injury such as Parkinson's disease,
CC Alzheimer's disease, Down's syndrome, stroke, cardiac arrest, sciatic
CC crush, spinal cord injury, injury to sensory neurons or degenerative
CC disease of the retina. The present sequence represents the murine Mts 1
CC protein, from which the Mts proteins of the invention may be derived.
CC NOTE: The present sequence is not shown in the specification but it has
CC been derived from the sequence given in Genes Dev. 3, 1989 (referred to
CC on page 10 of the specification)
XX
XX Sequence 101 AA;
SQ
Query Match 96.2%; Score 77; DB 4; Length 101;
Best Local Similarity 93.8%; Pred. No. 2.7e-05;
Matches 15; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 KEGDKFKLNKSELKEL 16
DB 22 KEGDKFKLNKTELKEL 37
|||||:|||||
RESULT 12
ADE61764
ID ADE61764 standard; protein; 93 AA.
AC ADE61764;
XX
XX 29-JAN-2004 (first entry)
XX Human Protein P23297, SEQ ID NO 7691.
XX
XX Homo sapiens.
XX Human; pain; neuronal tissue; gene therapy;
XX spinal segmental nerve injury; chronic constriction injury; CCI;
XX spared nerve injury; SNI; Chung.
XX
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX 01-NOV-2001; 2001US-0346382P.
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX
XX Woolf C, D'urso D, Befort K, Costigan M;
XX
XX WPI; 2003-268312/26.
XX GENBANK; P23297.
XX
XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.
XX
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC the sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 93 AA;
SQ

Query Match 85.0%; Score 68; DB 7; Length 93;
Best Local Similarity 81.2%; Pred. No. 0.00084;
Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 KEGDKFKLNKSELKEL 16
DB 21 KEGDKYKLSKSELKEL 36
|||||:|||||

RESULT 13
AAB45531
ID AAB45531 standard; protein; 94 AA.
XX
XX AAB45531;
XX
XX 22-FEB-2001 (first entry)
XX Human S100A1 protein.
XX
XX S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;
XX calcium-binding protein; calcium homeostasis; cardiac muscle;
XX pumping capacity; myocardial cell; systolic calcium ion release;
XX sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;
XX valve defect.
XX
XX Homo sapiens.
XX
XX DE19915485-A1.
XX
XX 19-OCT-2000.
XX
XX 07-APR-1999; 99DE-01015485.
XX
XX 07-APR-1999; 99DE-01015485.
XX (KATU/) KATUS H A.
XX (REMP/) REMPEIS A.
XX
XX Katus HA, Remppis A;
PI

XX WPI; 2000-673510/66.
 DR N-PSDB; AAC81801.
 XX
 XX Composition containing S100 protein, corresponding nucleic acid or
 PT vector, useful for treating cardiomyopathy and cardiac insufficiency.
 PT
 XX Claim 10; Page 8; 36pp; German.
 PS
 XX This invention describes a novel composition for treating primary or
 CC secondary cardiomyopathy or cardiac insufficiency contains at least one
 CC S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or
 CC fragments, or a gene transfer vector containing (II), optionally
 CC formulated with auxiliaries and/or carriers. (I) are calcium-binding
 CC proteins involved in calcium homeostasis, so their overexpression in
 CC cardiac muscle will improve pumping capacity (and overall capacity) of
 CC the heart. In cultured myocardial cells they increase the contraction and
 CC relaxation rates associated with increased systolic calcium ion release
 CC from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are
 CC used to treat cardiomyopathy (CMP) where inherited or caused by
 CC spontaneous mutations and ischemic CMP caused by arteriosclerosis,
 CC dilative CMP caused by toxic/infectious disease, cardiac disease caused
 CC by pulmonary and/or arterial hypertension, and structural disease caused
 CC by rhythm disorders or valve defects, generally any condition associated
 CC with reduced contractile force. Unlike calmodulin, which is expressed
 CC ubiquitously, (I) show tissue-specific expression and treat the
 CC underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac
 CC disease
 XX
 XX Sequence 94 AA;
 SQ
 Query Match 85.0%; Score 68; DB 3; Length 94;
 Best Local Similarity 81.2%; Pred. No. 0.00085;
 Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 1 KEGDKFKLNKSELKEL 16
 Db 22 KEGDKYKLSKKELKEL 37
 RESULT 14
 AAM40258
 ID AAM40258 standard; protein; 94 AA.
 XX
 AC AAM40258;
 XX
 XX 22-OCT-2001 (first entry)
 DT
 XX
 DE Human polypeptide SEQ ID NO 3403.
 XX
 KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX
 OS Homo sapiens.
 XX
 XX W0200153312-A1.
 PN
 XX 26-JUL-2001.
 PD
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 XX 23-DEC-1999; 99US-00471275.
 PR
 PR 21-JAN-2000; 2000US-00488725.
 PR
 PR 25-APR-2000; 2000US-00552317.
 PR
 PR 20-JUN-2000; 2000US-00598042.
 PR
 PR 19-JUL-2000; 2000US-00620312.
 PR
 PR 03-AUG-2000; 2000US-00653450.
 PR
 PR 14-SEP-2000; 2000US-00662191.
 PR
 PR 19-OCT-2000; 2000US-00693036.

PR 29-NOV-2000; 2000US-00727344.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX
 XX WPI; 2001-442253/47.
 DR
 DR N-PSDB; AAI59414.
 XX
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 PT
 XX Example 5; SEQ ID NO 3403; 10078pp; English.
 PS
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX
 XX Sequence 94 AA;
 SQ
 Query Match 85.0%; Score 68; DB 4; Length 94;
 Best Local Similarity 81.2%; Pred. No. 0.00085;
 Matches 13; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 OY 1 KEGDKFKLNKSELKEL 16
 Db 22 KEGDKYKLSKKELKEL 37
 RESULT 15
 ABG96406
 ID ABG96406 standard; protein; 94 AA.
 XX
 AC ABG96406;
 XX
 XX 11-DEC-2002 (first entry)
 DT
 XX
 DE Human ovarian cancer marker OV55.
 XX
 KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
 KW central nervous system disorder; bacterial meningitis; viral meningitis;
 KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
 KW brain herniation; inflammation; encephalitis; testicular disorder;
 KW nontuberculous granulomatous orchitis; connective tissue disorder;
 KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
 KW histological type; carcinogenic; ovarian cancer marker.
 XX
 OS Homo sapiens.
 XX
 XX W0200271928-A2.
 PN
 XX 19-SEP-2002.
 PD
 XX
 PF 14-MAR-2002; 2002WO-US007826.
 XX
 XX 14-MAR-2001; 2001US-0276025P.
 PR
 PR 14-MAR-2001; 2001US-0276026P.
 PR
 PR 10-AUG-2001; 2001US-0311732P.
 PR
 PR 19-SEP-2001; 2001US-0323580P.

26-SEP-2001; 2001US-0324967P.
PR
26-SEP-2001; 2001US-0325102P.
PR
26-SEP-2001; 2001US-0325149P.
PR
XX
(WILL-) MILLENNIUM PHARM INC.
PA
XX Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
XX Meyers RE, Morrisey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
PI East RC, Lu K, Schmandt RE, Zhao X, Glatt K;
PI
XX WPI; 2002-723277/78.
XX
XX N-PSDB; ABS76505.
DR
XX
XX Assessing whether a patient is afflicted with ovarian cancer, useful in
PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
PT from a non cancer patient.
XX
XX Disclosure; Page 391; 481pp; English.
XX
XX The present invention relates to a new method for assessing whether a
CC patient is afflicted with ovarian cancer. The method involves comparing
CC the expression level of a marker in a patient sample and the normal level
CC of expression of the marker in a control non-ovarian cancer sample, where
CC the marker is selected from 363 cancer markers described in the
CC specification. The method of the invention is useful in diagnosing or
CC characterizing cancer, in detecting the presence of cancer as early as
CC possible, and the recurrence of ovarian cancer. The method may also be of
CC particular use with patients having an enhanced risk of developing
CC ovarian cancer (e.g. patients having a familial history of ovarian
CC cancer). The cancer markers may be used in the management and treatment
CC of e.g. brain and central nervous system disorders (e.g. bacterial and
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC testicular disorders (e.g. nontuberculous granulomatous orchitis),
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC disease or atherosclerosis). The compositions and methods may also be
CC used in assessing the histological type of neoplasm associated with
CC ovarian cancer, monitoring the progression of ovarian cancer, determining
CC whether ovarian cancer has metastasized or is likely to metastasize,
CC selecting a composition for inhibiting ovarian cancer, assessing the
CC ovarian carcinogenic potential of a compound, or inhibiting ovarian
CC cancer or at risk of developing ovarian cancer. The present amino acid
CC sequence represents one of the ovarian cancer markers described in the
CC invention
XX
XX Sequence 94 AA:
XX

Query Match	85.0%;	Score 68;	DB 5;	Length 94;
Best Local Similarity	81.2%;	Pred. No. 0.00085;		
Mismatches	2;	Mismatches	1;	Indels
Conservative				
Gaps	0;			

Qy 1 KEGDKFKLNKSELKEL 16
 |||||:||||
nb 22 KEGDKYKLKKELKEL 37

Search completed: September 9, 2004, 10:53:04
Job time : 34.6516 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:42:57 ; Search time 16.6452 Seconds
(without alignments)
189.556 Million cell updates/sec

Title: US-10-067-618-4
Perfect score: 52
Sequence: 1 ACPLKALDV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_xvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	43	82.7	779	4 Q9H348	Q9h348 homo sapien
2	43	82.7	849	11 Q8CG01	Q8CG01 mus musculus
3	43	82.7	852	11 Q8OV99	Q8OV99 mus musculus
4	43	82.7	853	11 Q8CJJC	Q8CJC3 mus musculus
5	43	82.7	853	11 Q8BKX6	Q8bkx6 mus musculus
6	43	82.7	854	4 Q8GTP8	Q8GTP8 homo sapien
7	39	75.0	101	6 Q9TV56	Q9TV56 canis famli
8	39	75.0	539	5 Q18680	Q18680 drosophila
9	39	75.0	539	5 Q874G6	Q874G6 drosophila
10	38	73.1	94	16 Q89Z43	Q89Z43 bacteroides
11	38	73.1	377	11 Q9D9L5	Q9D9L5 mus musculus
12	38	73.1	381	11 Q9CXH7	Q9CXH7 mus musculus
13	38	73.1	423	5 Q8IIB7	Q8IIB7 plasmodium
14	38	73.1	570	5 Q9XYAL	Q9XYAL limulus pol
15	37	71.2	600	13 Q8JI21	Q8JI21 oryzias lat
16	37	71.2	634	13 Q9W646	Q9W646 oryzias lat

17	36	69.2	112	16 Q8IUM8	Q8IUM8 bacillus an
18	36	69.2	112	16 Q8IHF8	Q8IHF8 bacillus ce
19	36	69.2	126	16 Q92DB3	Q92DB3 listeria in
20	36	69.2	126	16 Q8Y8J8	Q8Y8J8 listeria mo
21	36	69.2	167	17 Q8U107	Q8U107 pyrococcus
22	36	69.2	231	5 Q9XUX1	Q9XUX1 caenorhabdi
23	36	69.2	296	16 Q8CP49	Q8CP49 staphylococ
24	36	69.2	389	13 Q98T20	Q98T20 brachydanio
25	36	69.2	391	13 Q9DFH3	Q9DFH3 gallus gall
26	36	69.2	398	13 Q8AWD8	Q8AWD8 brachydanio
27	36	69.2	442	3 Q74677	Q74677 pichia angu
28	36	69.2	449	5 Q9VG86	Q9VG86 drosophila
29	36	69.2	660	16 Q8XAS5	Q8XAS5 leptospira
30	36	69.2	1051	10 Q9LXN4	Q9LXN4 arabidopsis
31	35	67.3	118	16 Q9KEW2	Q9KEW2 bacillus ha
32	35	67.3	127	17 Q8TTU2	Q8TTU2 methanosarc
33	35	67.3	147	16 Q8ZNJ2	Q8ZNJ2 salmonella
34	35	67.3	187	10 Q7XUP6	Q7XUP6 oryza sativ
35	35	67.3	209	4 Q96LL5	Q96LL5 homo sapien
36	35	67.3	288	16 Q99TY7	Q99TY7 staphylococ
37	35	67.3	288	16 Q8NWE5	Q8NWE5 staphylococ
38	35	67.3	300	17 Q97YV0	Q97YV0 sulfolobus
39	35	67.3	347	4 Q8N993	Q8N993 homo sapien
40	35	67.3	402	17 Q58335	Q58335 pyrococcus
41	35	67.3	482	4 Q60428	Q60428 homo sapien
42	35	67.3	538	2 Q84FL3	Q84FL3 pantoaea agg
43	35	67.3	578	16 Q8DCK1	Q8DCK1 vibrio vuln
44	35	67.3	578	16 Q87L91	Q87L91 vibrio para
45	35	67.3	818	16 Q9I580	Q9I580 pseudomonas

ALIGNMENTS

RESULT 1
Q9H348 ID Q9H348 PRELIMINARY; PRT; 779 AA.
AC Q9H348; (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE HVPS41p.
DE HVPS41p.
GN HVPS41.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aorta;
RA Zhao Y., Cao H., Jiang Y., Meng X., Zhao X., Liu D., Ding J.;
RT "Cloning and characterization of an alternative splice variant of
RT HVSP41 gene in human aorta."
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF135593; AAG43279.1; .
DR InterPro; IPR000547; Clathrin_repeat.
DR InterPro; IPR001841; Znf ring.
DR Pfam; PF00637; Clathrin; 1.
DR SMART; SM00299; CLH; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS50389; ZF_RING_2; 1.
SQ SEQUENCE 779 AA; 90333 MW; B75D88ACC3F3145D CRC64;

Query Match 82.7%; Score 43; DB 4; Length 779;
Best Local Similarity 77.8%; Pred. No. 4.6;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 CPLEXALDV 10

Db 563 CPLEXALEI 571

RESULT 2

```
Q8CG01 Q8CG01 PRELIMINARY; PRT; 849 AA.
AC Q8CG01;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Similar to vacuolar protein sorting 41 (Yeast) (Fragment).
GN VPS41.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Kidney;
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC023243; AAH23243.1; -.
DR MGD; MGI:1929215; Vps41.
DR InterPro; IPR000547; Clathrin_repeat.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00637; Clathrin; 1.
DR SMART; SM00299; CLH; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
FT NON_TER
SQ SEQUENCE 849 AA; 98199 MW; 33BE793D86D72C7F CRC64;

Query Match 82.7%; Score 43; DB 11; Length 849;
Best Local Similarity 77.8%; Pred. No. 5;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
DB 633 CPLEKALEI 641

RESULT 3
Q8OV99 Q8OV99 PRELIMINARY; PRT; 852 AA.
AC Q8OV99;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 25, Last annotation update)
DE Similar to vacuolar protein sorting 41 (Yeast) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CZECH II; TISSUE=Breast tumor;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC049916; AAH49916.1; -.
DR InterPro; IPR000547; Clathrin_repeat.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00637; Clathrin; 1.
DR SMART; SM00299; CLH; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
FT NON_TER
SQ SEQUENCE 852 AA; 98470 MW; 521AF8B1A3071088 CRC64;

Query Match 82.7%; Score 43; DB 11; Length 852;
Best Local Similarity 77.8%; Pred. No. 5;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
DB 636 CPLEKALEI 644

RESULT 4
Q8CJC3 Q8CJC3 PRELIMINARY; PRT; 853 AA.
AC Q8CJC3;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Vacuolar protein sorting 41.
GN VPS41.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/HeJ; TISSUE=Liver;
RA Dell'Angelica E.C.; Peters L.L.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF327407; AAH73032.1; -.
DR MGD; MGI:1929215; Vps41.
DR InterPro; IPR000547; Clathrin_repeat.
DR InterPro; IPR008941; TPR-like.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00637; Clathrin; 1.
DR SMART; SM00299; CLH; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
FT NON_TER
SQ SEQUENCE 853 AA; 98609 MW; 2C68F27B0CDE0539 CRC64;

Query Match 82.7%; Score 43; DB 11; Length 853;
Best Local Similarity 77.8%; Pred. No. 5;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
DB 637 CPLEKALEI 645

RESULT 5
Q8BKK6 Q8BKK6 PRELIMINARY; PRT; 853 AA.
AC Q8BKK6;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
DE Vacuolar assembly protein VPS41 homolog.
GN VPS41.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK051658; BAC34707.1; -.
DR MGD; MGI:1929215; Vps41.
DR InterPro; IPR000547; Clathrin_repeat.
DR InterPro; IPR001841; Znf_ring.
DR Pfam; PF00637; Clathrin; 1.
DR SMART; SM00299; CLH; 1.
DR PROSITE; PS50089; ZF_RING_2; 1.
FT NON_TER
SQ SEQUENCE 853 AA; 98647 MW; 3ABE9AE27015019E CRC64;

Query Match 82.7%; Score 43; DB 11; Length 853;
Best Local Similarity 77.8%; Pred. No. 5;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
DB 637 CPLEKALEI 645
```

Db 637 CPLEKALEI 645

RESULT 6

Q86TP8 PRELIMINARY; PRT; 854 AA.
 AC Q86TP8;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Similar to vacuolar protein sorting 41 (yeast).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC044851; AAH4851.1; -.
 DR InterPro; IPR000547; Clathrin_repeat.
 DR InterPro; IPR001841; Znf_ring.
 DR Pfam; PF00637; Clathrin; 1.
 DR SMART; SM00299; CLH; 1.
 DR PROSITE; PS00089; ZF_RING_2; 1.
 SQ SEQUENCE 854 AA; 58565 MW; 037577D5FA3A18A0 CRC64;

Query Match 82.7%; Score 43; DB 4; Length 854;
 Best Local Similarity 77.8%; Pred. No. 5;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CPLEKALDV 10

Db 638 CPLEKALEI 646

RESULT 7

Q9TV56 PRELIMINARY; PRT; 101 AA.
 AC Q9TV56;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE Metacasin.
 GN MTS1.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Madin-Darby; TISSUE=Kidney;
 RA Miyamori H., Hasegawa K., Kim K., Sato H.;
 RT "Expression of metastasis associated mts1 gene is co-induced with
 RT membrane type-1 matrix metalloproteinase (MT1-MMP) during oncogenic
 RT transformation and tubular formation of madin darby canine kidney
 RT (MDCK) epithelial cells."
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 CC 1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
 DR EMBL; AB031064; BAA83419.1; -.
 DR HSSP; P30801; 1A03.
 DR GO; GO:0005509; F:calcium ion binding; IEA.
 DR InterPro; IPR001751; CaBP_S100.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; efhand; 1.
 DR Pfam; PF01023; S_100; 1.
 DR ProDom; PD003407; CaBP_S100; 1.
 DR ProDom; PD000012; EF-hand; 1.
 DR PROSITE; PS00018; EF_HAND; 1.
 DR PROSITE; PS00303; S100_CaBP; 1.
 SQ SEQUENCE 101 AA; 11847 MW; 5535387DB7577DF0 CRC64;

Query Match 75.0%; Score 39; DB 6; Length 101;
 Best Local Similarity 100.0%; Pred. No. 4.3;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PLEKALDV 10

Db 4 PLEKALDV 11

RESULT 8

O18680 PRELIMINARY; PRT; 539 AA.
 ID O18680;
 AC O18680;
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
 DE ALAS protein.
 GN ALAS OR CG3017.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Shandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Doolittle R.F., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 "The genome sequence of Drosophila melanogaster."
 Science 287:2185-2195 (2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Ruiz de Mena I.R., Fernandez M.A., Bornstein B., Garesse R.;
 RA Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RA Celniker S.E., Aghavani A., Arcaina T.T., Baxter E., Blazej R.G.,

```
RA Chew M., Doyle C.M., Farfan D.E., Flanagan J., Houston K.A.,
RA Hummasti S.R., Karra K., Kearney L., Kim S.H., Lee B., Lomotan M.A.,
RA Mazda P., Mok M.S., Moshrefi A.R., Moshrefi M., Nixon K., Pacleb J.M.,
RA Park S., Pfeiffer B., Punch D., Snir E., Twomey B., Wan K.H.,
RA Whitelaw K.R., Yee A., Zhang R., Zieran L.L., Kimmel B.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003462; AAF47119.1; -.
DR EMBL; Y14577; CAA74915.1; -.
DR EMBL; Y14576; CAA74914.1; -.
DR EMBL; AC004642; -; NOT_ANNOTATED_CDS.
DR HSSP; P12998; 1BS0.
DR FlyBase; FBgn020764; Alas.
DR GO; GO:0003870; F:5-aminolevulinate synthase activity; IEA.
DR GO; GO:0008415; F:acyltransferase activity; IEA.
DR GO; GO:0008483; F:transaminase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biocytin synthesis; IEA.
DR GO; GO:0006783; P:heme biosynthesis; IEA.
DR InterPro; IPR003408; Ala_synthase.
DR InterPro; IPR004839; Aminotrans_I/II.
DR InterPro; IPR001917; Aminotrans_II.
DR Pfam; PF02490; Ala_synthase; 1.
DR Pfam; PF00155; aminotran 1.2; 1.
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; 1.
KW Acyltransferase; Transferase.
SQ SEQUENCE 539 AA; 58751 MW; FED46ADBD225D9A CRC64;

Query Match 75.0%; Score 39; DB 5; Length 539;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
DB 288 CPLEELLDV 296
|||||

RESULT 9
Q8T4G6 PRELIMINARY; PRT; 539 AA.
AC Q8T4G6;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SD01515P.
GN ALAS OR CG3017.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Murgall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY084198; AAL89936.1; -.
DR FlyBase; FBgn020764; Alas.
DR GO; GO:0003870; F:5-aminolevulinate synthase activity; IEA.
DR GO; GO:0008483; F:transaminase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biocytin synthesis; IEA.
DR GO; GO:0006783; P:heme biosynthesis; IEA.
DR InterPro; IPR003408; Ala_synthase.
DR InterPro; IPR004839; Aminotrans_I/II.
DR InterPro; IPR001917; Aminotrans_II.
DR Pfam; PF02490; Ala_synthase; 1.
DR Pfam; PF00155; aminotran 1.2; 1.
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; 1.
SQ SEQUENCE 539 AA; 58778 MW; 5D75ADC79CE541EF CRC64;
```

```
Query Match 75.0%; Score 39; DB 5; Length 539;
Best Local Similarity 77.8%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 0; Gaps 0;

QY 2 CPLEKALDV 10
DB 288 CPLEELLDV 296
|||||

RESULT 10
Q89Z43 PRELIMINARY; PRT; 94 AA.
AC Q89Z43;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN BT4534.
OS Bacteroides thetaiotaomicron.
OC Bacteriia; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AE016945; AAO79639.1; -.
DR GO; GO:0002888; C:flagellum (sensu Bacteria); IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0001539; P:ciliary/flagellar motility; IEA.
DR InterPro; IPR001029; Flagellin_C.
DR InterPro; IPR001387; HTH_3.
DR Pfam; PF01381; HTH_3; 1.
DR ProDom; PD000316; Flagellin_C; 1.
DR SMART; SM00530; HTH_XRE; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 94 AA; 10671 MW; 80CAA2573E17D133 CRC64;

Query Match 73.1%; Score 38; DB 16; Length 94;
Best Local Similarity 77.8%; Pred. No. 6.5;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
DB 76 CKIEKALDV 84
|||||

RESULT 11
Q9D9L5 PRELIMINARY; PRT; 377 AA.
AC Q9D9L5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE 1700052K15Rik protein (Hypothetical actin and actin-like containing
DE ARPM2) (Arpm2-pending protein).
GN ARPM2 OR 1700052K15RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Testis;
RC MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Iehii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
```

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staib F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustineich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Wittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.,
RA "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12468651;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Dratchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls U., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RX Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: ACTINS ARE HIGHLY CONSERVED PROTEINS THAT ARE INVOLVED
CC IN VARIOUS TYPES OF CELL MOTILITY AND ARE UBIQUITOUSLY EXPRESSED
CC IN ALL EUKARYOTIC CELLS (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE ACTIN FAMILY.
DR ENBL; AK006772; BAB24733.1; -.
DR ENBL; AK077010; BAC36563.1; -.
DR ENBL; BC052344; AAH52344.1; -.
DR MGD; MGI:1920603; Arpm2.
DR GO; GO:0015629; C:actin cytoskeleton; IEA.
DR GO; GO:0005200; F:structural constituent of cytoskeleton; IEA.
DR InterPro; IPR004000; Actin_like.
DR Pfam; PF00022; actin; 1.
DR PRINTS; PR00190; ACTIN.
DR SMART; SM00268; ACTIN; 1.
DR PROSITE; PS01132; ACTINS ACT LIKE; 1.
KW Hypothetical protein; Structural protein.
SQ SEQUENCE 377 AA; 41615 MW; BCDEF0F34BC5C88B6 CRC64;

Query Match 73.1%; Score 38; DB 11; Length 377;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2 CPLEKAL 8
Db 206 CPLEKAL 212
RESULT 12
ID Q9QXH7 PRELIMINARY; PRT; 381 AA.
AC Q9QXH7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Prostatic acid phosphatase.
GN ACP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Crew M.D., Chatta G.S., Borg C.D.;
RT "Sequence and expression of mouse prostatic acid phosphatase.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR ENBL; AF210243; AAF23171.1; -.
DR HSSP; P20646; IRPA.
DR MGD; MGI:1928480; Acp.
DR GO; GO:0003993; F:acid phosphatase activity; IEA.
DR InterPro; IPR000560; HisAc phosphatase.
DR Pfam; PF00328; acid phosphatase; 1.
DR PROSITE; PS00616; HIS ACID PHOSPHAT 1; 1.
DR PROSITE; PS00778; HIS ACID PHOSPHAT-2; 1.
SQ SEQUENCE 381 AA; 43689 MW; 60AD3919D77CB241 CRC64;
Query Match 73.1%; Score 38; DB 11; Length 381;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
Qy 1 ACPLEKALDV 10
Db 349 SCPLKXFDV 358
RESULT 13
ID Q8IIB7 PRELIMINARY; PRT; 423 AA.
AC Q8IIB7;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Ethanolamine kinase, putative.
GN PF11_0257.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=3D7;
RX MEDLINE=22255705; PubMed=12368864;
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eissen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairclimb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.;
RT "Genome sequence of the human malaria parasite Plasmodium

```

RT falciparum.";
RL Nature 419.498-511 (2002).
DR EMBL; AB014839; AAN35841.1; -.
DR GO; GO:0016301; F:kinase activity; IEA.
DR InterPro; IPR002573; Choline_kinase.
DR Pfam; PF01633; Choline_Kinase; 1.
KW Kinase.
SQ SEQUENCE 423 AA; 49914 MW; C57ADFE8121A992A CRC64;

Query Match 73.1%; Score 38; DB 5; Length 423;
Best Local Similarity 50.0%; Pred. No. 27;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACPELKALDV 10
Db 317 SCPMERAYDI 326

RESULT 14
Q9XYA1 PRELIMINARY; PRT; 570 AA.
AC Q9XYA1;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 5-aminolevulinate synthase (EC 2.3.1.37).
GN ALS.
OS Limulus polyphemus (Atlantic horseshoe crab).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC Limulidae; Limulus.
OX NCBI_TaxID=6850;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Hepatopancreas;
RX MEDLINE=99261649; PubMed=10331264;
RA Duncan R., Faggart M.A., Roger A.J., Cornell N.W.;
RT "Phylogenetic analysis of the 5-aminolevulinate synthase gene.";
RL Mol. Biol. Evol. 16:383-396(1999).
DR EMBL; AF073336; AAD20805.1; -.
DR HSSP; P12998; IBSO.
DR GO; GO:0003870; F:5-aminolevulinate synthase activity; IEA.
DR GO; GO:0008415; F:acyltransferase activity; IEA.
DR GO; GO:0008483; F:transaminase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0009058; P:biosynthesis; IEA.
DR GO; GO:0006783; P:heme biosynthesis; IEA.
DR InterPro; IPR003408; Ala_synthase.
DR InterPro; IPR004839; Aminotrans_I/II.
DR InterPro; IPR001917; Aminotrans_II.
DR Pfam; PF02490; ALA synthase; 1.
DR Pfam; PF00155; aminotran.1.2; 1.
DR PROSITE; PS00599; AA_TRANSFER_CLASS_2; 1.
KW Acyltransferase; Transferase.
SQ SEQUENCE 570 AA; 62998 MW; 521B3D13453F4C94 CRC64;

Query Match 73.1%; Score 38; DB 5; Length 570;
Best Local Similarity 77.8%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
Db 319 CPLEKMCVDV 327

RESULT 15
Q8JI21 PRELIMINARY; PRT; 600 AA.
AC Q8JI21;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Chorlogenin Hminor (Fragment).
OS Oryzias latipes (Medaka fish) (Japanese ricefish).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HNI;
RA Kanamori A., Hori H.;
RT "Genomic organization of ZP domain containing egg envelope genes in medaka.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF396669; FAN91821.1; -.
DR GO; GO:0045735; F:nutrient reservoir activity; IEA.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR001419; Glutenin.
DR InterPro; IPR000519; P_trefoil.
DR Pfam; PF00088; trefoil; 1.
DR Pfam; PF00100; zona_pellucida; 1.
DR PRINTS; PR00210; GLUTENIN.
DR SMART; SM00018; PD; 1.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS00025; P_TREFOIL; 1.
FT NON_TER 1
SQ SEQUENCE 600 AA; 67416 MW; 145169D26D679DDC CRC64;

Query Match 71.2%; Score 37; DB 13; Length 600;
Best Local Similarity 77.8%; Pred. No. 61;
Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
Db 565 CPLRKARDV 573

Search completed: September 9, 2004, 10:58:16
Job time : 19.6452 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:32:57 ; Search time 2.83871 Seconds
(without alignments)
183.429 Million cell updates/sec

Title: US-10-067-618-4
Perfect score: 52
Sequence: 1 ACPLKALDV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	100.0	101	1	S104_HUMAN
2	43	82.7	854	1	VP41_HUMAN
3	41	78.8	100	1	S104_BOVIN
4	38	73.1	340	1	TRPD_BACPU
5	38	73.1	381	1	GSFP_KLEPN
6	36	69.2	101	1	S104_MOUSE
7	36	69.2	101	1	S104_RAT
8	36	69.2	766	1	METE_PSEAE
9	35	67.3	429	1	SYH_STRRG
10	34	65.4	92	1	S106_HORSE
11	34	65.4	128	1	YDEP_BACSU
12	34	65.4	243	1	UBIE_RALSO
13	34	65.4	336	1	HXDA_HETER
14	34	65.4	402	1	ILBC_MOUSE
15	34	65.4	404	1	GSFP_ERWCH
16	34	65.4	438	1	VATB_CHLCV
17	34	65.4	438	1	VATB_CHLPU
18	34	65.4	438	1	VATB_CHLPN
19	34	65.4	438	1	VATB_CHLTR
20	34	65.4	582	1	HEMO_DELLE
21	34	65.4	582	1	HEMO_OPSTA
22	34	65.4	583	1	HEMO_BRARE
23	34	65.4	587	1	HEMO_HUMAN
24	34	65.4	587	1	HEMO_MOUSE
25	34	65.4	587	1	HEMO_RAT
26	34	65.4	635	1	HEM1_CHICK
27	34	65.4	640	1	HEM1_DELLE
28	34	65.4	640	1	HEM1_HUMAN
29	34	65.4	642	1	HEM1_RAT
30	34	65.4	1294	1	YA3B_SCHPO
31	33	63.5	85	1	RNPH_RHOCA
32	33	63.5	89	1	S106_MOUSE
33	33	63.5	90	1	S106_HUMAN

34	33	63.5	90	1	S106_RAT
35	33	63.5	114	1	RSN_MOUSE
36	33	63.5	125	1	Y1BR_BACSU
37	33	63.5	155	1	Y115_YETJA
38	33	63.5	183	1	YF36_PYRFU
39	33	63.5	263	1	KKA9_STRRI
40	33	63.5	299	1	NK25_XENLA
41	33	63.5	330	1	DUFF_BOVIN
42	33	63.5	357	1	YE28_HELPY
43	33	63.5	520	1	LAC1_TRAHI
44	33	63.5	894	1	WP3A_BACSU
45	32	61.5	202	1	TRPF_LISMO

ALIGNMENTS

RESULT 1					
S104_HUMAN					
ID	S104_HUMAN	STANDARD;	PRT;	101 AA.	
AC	P26447;				
DT	01-AUG-1992 (Rel. 23, Created)				
DT	01-AUG-1992 (Rel. 23, Last sequence update)				
DT	15-MAR-2004 (Rel. 43, Last annotation update)				
DE	Placental calcium-binding protein (Calvasculin) (S100 calcium-binding protein A4) (MTS1 protein).				
DE	protein A4) (MTS1 protein).				
GN	S100A4 OR CAPL OR MTS1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	SEQUENCE FROM N.A.				
RP	TISSUE=Heart;				
RX	MEDLINE=93041710; PubMed=1384693;				
RA	Engelkamp D., Schaefer B.W., Erne P., Heizmann C.W.;				
RT	"S100 alpha, CAPL, and CACY: molecular cloning and expression analysis of three calcium-binding proteins from human heart."				
RL	Biochemistry 31:10258-10264(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=93342039; PubMed=8341667;				
RA	Engelkamp D., Schaefer B., Mattei M.-G., Erne P., Heizmann C.W.;				
RT	"Six S100 genes are clustered on human chromosome 1q21: identification of two genes coding for the two previously unreported calcium-binding proteins S100D and S100E."				
RL	Proc. Natl. Acad. Sci. U.S.A. 90:6547-6551(1993).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Spleen;				
RA	MEDLINE=93028421; PubMed=1329089;				
RT	Tulchinsky E.M., Ford H.L., Kramerov D., Reshetnyak E., Grigorian M., Zain S., Lukanidin E.;				
RT	"Transcriptional analysis of the mts1 gene with specific reference to 5' flanking sequences."				
RL	Proc. Natl. Acad. Sci. U.S.A. 89:9146-9150(1992).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Cervix, and Prostate;				
RX	MEDLINE=22389257; PubMed=12479932;				
RA	Strausberg K.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler C., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Placental calcium-binding protein homolog.
GN S100A4 OR CAPL.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Retina;
RX MEDLINE=94164991; PubMed=81119967;
RA Polans A.S., Palczewski K., Asson-Batres M.A., Ohguro H., Witowska D.,
RA Halsey J.L., Balzer L., Crabb J.W.;
RT "Purification and primary structure of Capl, an S-100-related
calcium-binding protein isolated from bovine retina";
RL J. Biol. Chem. 269:6233-6240 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98430466; PubMed=9759666;
RA Duarte W.R., Kasugai S., Iimura T., Oida S., Takenaga K., Ohya K.,
RA Ishikawa I.;
RT "cDNA cloning of S100 calcium-binding proteins from bovine
periosteal ligament and their expression in oral tissues.";
RL J. Dent. Res. 77:1694-1699 (1998).
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: D89056; BAA13754.1; -.
CC PIR: A53217; A53217.
CC HSP: P30801; IA03.
CC InterPro: IPR001751; CAPS S100.
CC DR InterPro; IPR002048; EF-hand.
CC Pfam: PF00036; ehand; 1.
CC Pfam; PF01023; S100; 1.
CC ProDom; PD003407; CAPS S100; 1.
CC ProDom; PD000012; EF-hand; 1.
CC DR PROSITE; PS00018; EF-HAND; 1.
CC DR PROSITE; PS00303; S100 CAPB; 1.
CC KW Calcium-binding; Acetylation.
CC INIT MET 0
CC MOD_RES 0
CC FT CA_BIND 1 1 ACETYLATION.
CC FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
CC FT CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
CC SQ SEQUENCE 100 AA; 11675 MW; DFFCA7561D5EFA4 CRC64;

Query Match 78.8%; Score 41; DB 1; Length 100;
Best Local Similarity 90.0%; Pred. No. 0.45;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACPLKALDV 10
| | | | |
DB 1 AYPLEKALDV 10

RESULT 4
TRPD BACPU STANDARD; PRT; 340 AA.
AC P18261;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Anthranilate phosphoribosyltransferase (EC 2.4.2.18).
GN TRPD.
OS Bacillus pumilus (Bacillus mesentericus).

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1408;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RUB502;
RX MEDLINE=90236301; PubMed=21110100;
RA Rivas M.V., Jarvis E.D., Rudner R.;
RT "The structure of the trpE, trpD and 5' trpC genes of Bacillus
pumilus.";
RL Gene 87:71-78 (1990).
RN [2]
RP ERRATUM
RX MEDLINE=91033058; PubMed=2227447;
RA Rivas M.V., Jarvis E.D., Rudner R.;
RL Gene 94:141-143 (1990).
CC -!- CATALYTIC ACTIVITY: Anthranilate + phosphoribosyl-diphosphate =
CC N-5'-phosphoribosyl-anthranilate + diphosphate.
CC -!- PATHWAY: Tryptophan biosynthesis; second step.
CC -!- SIMILARITY: Belongs to the anthranilate phosphoribosyltransferase
CC family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M36468; AAB02273.1; -.
CC PIR: JH0099; JH0099.
CC HAMAP: MF_00211; -; 1.
CC InterPro; IPR005940; Ant_phospho trans.
CC DR InterPro; IPR000312; Glyco trans 3.
CC Pfam; PF02885; Glycos trans 3N; 1.
CC Pfam; PF00591; Glycos_transf_3; 1.
CC ProDom; PD001864; Glyco trans 3; 1.
CC DR TIGRFAWS; TIGR01245; trpD; 1.
CC KW Tryptophan biosynthesis; Transferase; Glycosyltransferase.
CC SQ SEQUENCE 340 AA; 36789 MW; 954230324ED77E7D CRC64;

Query Match 73.1%; Score 38; DB 1; Length 340;
Best Local Similarity 70.0%; Pred. No. 6.1;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACPLKALDV 10
| | | | |
DB 64 AAPMERALDV 73

RESULT 5
GSPF KLEPN STANDARD; PRT; 381 AA.
AC P15745;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE General secretion pathway protein F (Pullulanase secretion protein
DE pulF).
DE pulF.
OS Klebsiella pneumoniae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Klebsiella.
OX NCBI_TaxID=573;
RN [1]
RP SEQUENCE FROM N.A.
RA Pugsley A.P., D'Enfert C., Reyss I., Wandersman C.;
RL Submitted (MAR-1990) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INVOLVED IN A GENERAL SECRETION PATHWAY (GSP) FOR THE
CC EXPORT OF PROTEINS. REQUIRED FOR THE TRANSLOCATION OF PULLULANASE.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- SIMILARITY: BELONGS TO THE PULF/OUTF/EXEF/XPSF/XCPS FAMILY.

```
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sb-sib.ch).
-----
EMBL; M32613; AAA25128.1; -
PIR; S20035; S20035.
InterPro; IPR003004; Bac_GSPF.
InterPro; IPR001992; Bact_sec_systII.
Pfam; PF00482; GSP1 F; 1.
PRINTS; PR00812; BCTERIALGSPF.
PROSITE; PS00874; T2SP F; 1.
Transport; Transmembrane; Inner membrane.
TRANSMEM 145 165 POTENTIAL.
TRANSMEM 198 218 POTENTIAL.
TRANSMEM 353 373 POTENTIAL.
SEQUENCE 381 AA; 41872 MW; 01347C0994B29933 CRC64;
Query Match 73.1%; Score 38; DB 1; Length 381;
Best Local Similarity 88.9%; Pred. No. 6.8; 1; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 1;
QY 1 ACPLKALD 9
Db 77 AIPLEKALD 85
RESULT 6
S104 MOUSE
ID S104 MOUSE STANDARD; PRT; 101 AA.
AC P07091; P20066; -
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Placental calcium-binding protein (18A2) (PEL98) (MTS1 protein)
DE (Metastatic cell protein).
GN S100A4 OR CAPL OR MTS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
[1]
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/c;
RX MEDLINE=87316927; PubMed=3628004;
RA Jackson-Grusby L.L., Swiergiel J., Linzer D.I.H.;
RT "A growth-related mRNA in cultured mouse cells encodes a placental
calcium binding protein."
RL Nucleic Acids Res. 15:6677-6690 (1987).
[2]
RN 2
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/c;
RX MEDLINE=88198109; PubMed=3162911;
RA Goto K., Endo H., Fujiyoshi T.;
RT "Cloning of the sequences expressed abundantly in established cell
lines: identification of a cDNA clone highly homologous to S-100, a
calcium binding protein."
RL J. Biochem. 103:48-53 (1988).
[3]
RN 3
RP SEQUENCE FROM N.A.
RX MEDLINE=89378739; PubMed=2550322;
RA Ebralidze A., Tulchinsky E., Grigorian M., Afanasyeva A., Senin V.,
RA Revazova E., Lukanidin E.;
RT "Isolation and characterization of a gene specifically expressed in
different metastatic cells and whose deduced gene product has a high
degree of homology to a Ca2+-binding protein family."
RL Genes Dev. 3:1086-1093 (1989).
[4]
RN 4
RP SEQUENCE FROM N.A.
TISSUE=Liver;
MEDLINE=90236313; PubMed=2332170;
Tulchinsky E.M., Grigorian M.S., Ebralidze A.K., Milshina N.I.,
Lukanidin E.M.;
"Structure of gene mts1, transcribed in metastatic mouse tumor
cells."
RL Gene 87:219-223 (1990).
[5]
RN 5
RP SEQUENCE OF 1-54 FROM N.A.
MEDLINE=93141279; PubMed=8423998;
Tulchinsky E., Kramerov D., Ford H.L., Reshetnyak E., Lukanidin E.,
Zain S.;
"Characterization of a positive regulatory element in the mts1 gene."
RL Oncogene 8:79-86 (1993).
CC -!- TISSUE SPECIFICITY: Specifically expressed in different metastatic
cells.
CC -!- INDUCTION: The mRNA coding for this protein increases in
abundance after serum stimulation of quiescent mouse fibroblasts.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sb-sib.ch).
-----
EMBL; X05835; CAA29282.1; -
EMBL; D00208; BAA00148.1; -
EMBL; M36578; AAA39749.1; -
EMBL; M36579; AAA39750.1; -
EMBL; X16190; CAA34316.1; -
EMBL; X16094; CAA34224.1; -
PIR; S06207; S06207.
HSP; P30801; LA03.
MGD; MGI:1330282; S100a4.
InterPro; IPR001751; CaBP_S100.
InterPro; IPR002048; EF-hand.
Pfam; PF00036; ehand; 1.
ProDom; PD003407; CaBP_S100; 1.
PROSITE; PS00018; EF HAND; 1.
PROSITE; PS00303; S100 CABP; 1.
Calcium-binding; Placenta.
CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT CONFLICT 47 47 G -> GVSGSXFNQ (IN REF. 5).
FT CONFLICT 47 47 G -> GVSGSXFNQ (IN REF. 5).
SQ SEQUENCE 101 AA; 11721 MW; 2302254B67A4C873 CRC64;
Query Match 69.2%; Score 36; DB 1; Length 101;
Best Local Similarity 80.0%; Pred. No. 4.4;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 ACPLKALDV 10
Db 2 ARPLEEALDV 11
RESULT 7
S104 RAT
ID S104 RAT STANDARD; PRT; 101 AA.
AC P05942;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Placental calcium-binding protein (Nerve growth factor induced protein
42A) (p9K).
GN S100A4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC
```

```

OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88125019; PubMed=3422491;
RA Masiakowski P., Shooter E.M.;
RT "Nerve growth factor induces the genes for two proteins related to a
RT family of calcium-binding proteins in PC12 cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:1277-1281 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88118907; PubMed=3430604;
RA Barracough R., Savin J., Dube S.K., Rudland P.S.;
RT "Molecular cloning and sequence of the gene for p9Ka. A cultured
RT myoepithelial cell protein with strong homology to S-100, a calcium-
RT binding protein.";
RL J. Mol. Biol. 198:13-20 (1987).
RN [3]
RP SEQUENCE OF 3-101 FROM N.A.
RX MEDLINE=92158347; PubMed=1741158;
RA De Voe M.W., Mukherjee B.B.;
RT "Transformation of normal rat kidney cells by v-K-ras enhances
RT expression of transin 2 and an S-100-related calcium-binding
RT protein.";
RL Oncogene 7:109-119 (1992).
CC -!- INDUCTION: By nerve growth factor.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X06916; CAA30014.1; -
CC DR EMBL; J03628; AAA42098.1; -
CC DR EMBL; X64022; -; NOT_ANNOTATED_CDS.
CC DR EMBL; X64023; -; NOT_ANNOTATED_CDS.
CC DR PIR; S01759; S01759.
CC DR HSP; P30801; IA03.
CC DR InterPro; IPR001751; CaBP_S100.
CC DR InterPro; IPR002048; EF-hand.
CC DR Pfam; PF00036; ehand; 1.
CC DR Pfam; PF01023; S100; 1.
CC DR ProDom; PD003407; CaBP_S100; 1.
CC DR PROSITE; PS00018; EF_HAND; 1.
CC DR PROSITE; PS00303; S100_CABP; 1.
CC DR Calcium-binding.
CC KW CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
CC FT CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
CC SQ SEQUENCE 101 AA; 11776 MW; EA0619CEE4F487C1 CRC64;

Query Match 69.2%; Score 36; DB 1; Length 101;
Best Local Similarity 80.0%; Pred. No. 4.4;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACPLKALDV 10
Db 2 ARPLEALDV 11

RESULT 8
METE_PSEAE STANDARD; PRT; 766 AA.
AC F57703; Q91Z17;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 5-methyltetrahydropteroyltyr-glutamate--homocysteine methyltransferase
DE (EC 2.1.1.14) (Methionine synthase, vitamin-B12 independent isozyme)
DE (Cobalamin-independent methionine synthase).

```

```

GN METE OR PA1927.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
CC -!- FUNCTION: Catalyzes the transfer of a methyl group from 5-
CC methyltetrahydrofolate to homocysteine resulting in methionine
CC formation (By similarity).
CC -!- CATALYTIC ACTIVITY: 5-methyltetrahydropteroyl-L-glutamate + L-
CC homocysteine = tetrahydropteroyl-L-glutamate + L-methionine.
CC -!- COFACTOR: Zinc; binds one ion per subunit (By similarity).
CC -!- PATHWAY: Terminal step in the de novo biosynthesis of methionine.
CC -!- SIMILARITY: Belongs to the vitamin-B12 independent methionine
CC synthase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE004619; AAG05315.1; -
CC DR PIR; D83404; D83404.
CC DR HAMAP; MF 00172; -; 1.
CC DR InterPro; IPR006276; Met_syn_B12ind.
CC DR InterPro; IPR002629; Methionine_synth.
CC DR Pfam; PF01717; Methionine_synth.
CC DR ProDom; PD004692; Methionine_synth; 2.
CC DR TIGRFAMs; TIGR01371; met_syn_B12ind; 1.
CC KW Transferase; Methyltransferase; Methionine biosynthesis; Zinc; Repeat;
CC Complete proteome.
CC FT METAL 650 650 ZINC (BY SIMILARITY).
CC FT METAL 652 652 ZINC (BY SIMILARITY).
CC FT METAL 735 735 ZINC (BY SIMILARITY).
CC SQ SEQUENCE 766 AA; 86210 MW; 7D638917B844AB81 CRC64;

Query Match 69.2%; Score 36; DB 1; Length 766;
Best Local Similarity 77.8%; Pred. No. 34;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
Db 302 CDLEKALEV 310

RESULT 9
SYH_STRR6 STANDARD; PRT; 429 AA.
AC Q8DN46;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Histidyl-tRNA synthetase (EC 6.1.1.21) (Histidine--tRNA ligase)
DE (HISRS).
GN HISS OR SPR1931.
OS Streptococcus pneumoniae (strain ATCC BAA-255 / R6).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.

```


RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Hache J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Scrokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Taseuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vanier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*.";
 RL Nature 390:249-256(1997).
 CC -!- SIMILARITY: Belongs to the UPF0087 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; AB001488; BAA19363.1; ALT_INIT.
 CC EMBL; Z99106; CAB12336.1; -.
 CC PIR; B69779; B69779.
 CC Subtilist; BG12143; ydep.
 CC InterPro; IPR002577; DUF24.
 CC Pfam; PF01638; DUF24; 1.
 CC ProDom; PD004032; DUF24; 1.
 CC KW Hypothetical protein; Complete proteome.
 CC SEQUENCE 128 AA; 15211 MW; A47D416545CCAC58 CRC64;
 Query Match 65.4%; Score 34; DB 1; Length 128;
 Best Local Similarity 55.6%; Pred. No. 14;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 CPLEKALDV 10
 |||
 Db 18 CPVETILDI 26

RESULT 12
 UBIE_RALSO STANDARD; PRT; 243 AA.
 AC Q8Y278;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Ubiquinone/menaquinone biosynthesis methyltransferase ubie
 (EC 2.1.1.-).
 GN UBIE OR RSC0458 OR RS04441.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 ON NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangot S.,
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
 RA Chandler M., Choise N., Claudel-Renard C., Cunac S., Demange N.,

RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
 RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
 RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*.";
 RL Nature 415:497-502(2002)
 CC -!- FUNCTION: Methyltransferase required for the conversion of
 CC dimethylmenaquinone (DMQH2) to menaquinone (MKH2) and the
 CC conversion of 2-polyphenyl-6-methoxy-1,4-benzoquinol (DDMQH2) to
 CC 2-polyphenyl-3-methyl-6-methoxy-1,4-benzoquinol (DDMQH2) (By
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + 2-polyphenyl-6-
 CC methoxy-1,4-benzoquinol = S-adenosyl-L-homocysteine + 2-
 CC polyphenyl-3-methyl-6-methoxy-1,4-benzoquinol.
 CC -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + demethylmenaquinol =
 CC S-adenosyl-L-homocysteine + menaquinol.
 CC -!- PATHWAY: Menaquinone biosynthesis; last step.
 CC -!- PATHWAY: Ubiquinone biosynthesis;
 CC -!- SIMILARITY: Belongs to the ubie family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; AL646059; CAD13986.1; -.
 CC HAMAP; MF 01813; - 1.
 CC InterPro; IPR001601; Methyltransf.
 CC InterPro; IPR000051; SAM bind.
 CC InterPro; IPR004034; Ubi/men Mettransf.
 CC InterPro; IPR004033; UbiE/COQ5 Metrf.
 CC Pfam; PF01209; Ubie methyltran; 1.
 CC PROSITE; PS01183; Ubie 1; 1.
 CC PROSITE; PS01184; Ubie 2; 1.
 CC KW Menaquinone biosynthesis; Ubiquinone biosynthesis; Transferase;
 CC Methyltransferase; Complete proteome.
 CC SEQUENCE 243 AA; 27075 MW; 2A337D9EA2F43375 CRC64;
 Query Match 65.4%; Score 34; DB 1; Length 243;
 Best Local Similarity 87.5%; Pred. No. 26;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 PLEKALDV 10
 |||
 Db 170 PLEKAYDV 177

RESULT 13
 HXDA_HETFR STANDARD; PRT; 336 AA.
 AC Q9IAL4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Homeobox protein Hox-D10.
 GN HOXD10.
 OS Heterodontus francisci (Horn shark).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 OC Elasmobranchii; Galeomorphii; Heterodontidae; Heterodontiformes;
 OC Heterodontidae; Heterodontus.
 ON NCBI_TaxID=7792;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20144096; PubMed=10677514;
 RA Kim C.B., Amemiya C., Bailey W., Kawasaki K., Mezey J., Miller W.,
 RA Minoshima S., Shimizu N., Wagner G., Ruddle F.;
 RT "Hox cluster genomics in the horn shark, *Heterodontus francisci*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1655-1660(2000).
 CC -!- FUNCTION: Sequence-specific transcription factor which is part of
 CC a developmental regulatory system that provides cells with
 CC specific positional identities on the anterior-posterior axis (By

```

CC similarity).
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: Belongs to the Abd-B homeobox family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF224263; AAF4634.1; -.
CC HSP; P02833; 9ANT.
CC TRANSFAC; T04484; -.
CC InterPro; IPR001356; Homeobox.
CC Pfam; PF00046; homeobox.1.
CC PRINTS; PR00024; HOMEBOX.
CC ProDom; PD00010; Homeobox; 1.
CC SMART; SM00389; HOX; 1.
CC PROSITE; PS00027; HOMEBOX_1; 1.
CC PROSITE; PS00071; HOMEBOX_2; 1.
CC Homeobox; DNA-binding; Developmental protein; Nuclear protein;
KW Transcription regulation.
FT DNA BIND 262 321 HOMEBOX.
SQ SEQUENCE 336 AA; 38019 MW; 9C38BBAA8F420424 CRC64;

Query Match 65.4%; Score 34; DB 1; Length 336;
Best Local Similarity 77.8%; Pred. No. 37;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 ACPLKALD 9
Db 157 AYPLEKSLD 165

RESULT 14
IIBC MOUSE STANDARD; PRT; 402 AA.
AC P29452;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Interleukin-1 beta convertase precursor (IL1-BC) (EC 3.4.22.36) (IL-1
DE beta converting enzyme) (ICE) (interleukin-1 beta converting enzyme)
DE (P45) (Caspase-1) (CASP-1).
DE CASP1 OR IL1BC
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93056487; PubMed=1431103;
RA Nett-Flordalisi M.A., Cerretti D.P., Berson D.R., Gilbert D.J.,
RA Jenkins N.A., Copeland N.G., Black R.A., Chaplin D.D.;
RA "Molecular cloning of the murine IL-1 beta converting enzyme cDNA.";
RA J. Immunol. 149:3254-3259(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=93189587; PubMed=8446594;
RA Molineaux S.M., Casano F.J., Rolando A.M., Peterson E.P., Limjuco G.,
RA Chin J., Griffin P.R., Calaycay J.R., Ding G.J.-F., Yamin T.-T.,
RA Palaya O.C., Luell S., Fletcher D., Miller D.K., Howard A.D.,
RA Thornberry N.A., Kostura M.J.;
RA "Interleukin 1 beta (IL-1 beta) processing in murine macrophages
RA requires a structurally conserved homologue of human IL-1 beta
RA converting enzyme.";
RA Proc. Natl. Acad. Sci. U.S.A. 90:1809-1813(1993).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=129;
RX MEDLINE=94307735; PubMed=8034321;

```

Casano F.J., Rolando A.M., Mudgett J.S., Molineaux S.M.;
 "The structure and complete nucleotide sequence of the murine gene
 encoding interleukin-1 beta converting enzyme (ICE).";
 Genomics 20:474-481(1994).

[4]
 SEQUENCE FROM N.A.
 STRAIN=FVB/N; TISSUE=Mammary gland;
 MEDLINE=22388257; PubMed=12477932;
 Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 Fahey J., Helton E., Kettman M., Madan A.C., Shevchenko Y., Bouffard G.G.,
 Whiting M., Madan A., Young A.C., Green E.D., Dickson M.C.,
 Blakesley R.W., Touchman J.W., Schmutz J., Myers R.M.,
 Rodriguez A.C., Grimwood J., Krzywinski M.I., Skalska U., Smailus D.E.,
 Butterfield Y.S.N., Schein J.E., Jones S.J.M., Marra M.A.,
 Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

[5]
 CHARACTERIZATION.
 STRAIN=C3H/An;
 RX MEDLINE=97190206; PubMed=9038361;
 van de Craen M., Vandenabeele P., Declercq W., van den Brande I.,
 van Loo G., Molemans F., Schotte P., van Crielinge W., Beyaert R.,
 Fiers W.;
 "Characterization of seven murine caspase family members.";
 FEBS Lett. 403:61-69(1997).

-1- FUNCTION: Involved in the activation cascade of caspases
 responsible for apoptosis execution. Cleaves and activates sterol
 regulatory element binding proteins (SREBPs). Overexpression
 promotes programmed cell death (By similarity).
 -1- FUNCTION: Thiol protease that cleaves IL-1 beta between an Asp and
 an Ala, releasing the mature cytokine which is involved in a
 variety of inflammatory processes.
 -1- CATALYTIC ACTIVITY: Release of interleukin 1-beta by specific
 cleavage at 116-Asp|-Ala-117 and 27-Asp|-Gly-28 bonds in
 precursor. Also hydrolyzes the small-molecule substrate, Ac-Tyr-
 Val-Ala-Asp|-NHMe.
 -1- SUBUNIT: Tetramer that consists of two heterodimer of a 20 kDa
 (p20) and a 10 kDa (p10) subunits.
 -1- SUBCELLULAR LOCATION: Cytoplasmic.
 -1- TISSUE SPECIFICITY: High level expression seen in spleen and
 lung, low level expression seen in brain, heart, liver, kidney,
 testis and skeletal muscle.
 -1- PTM: The two subunits are derived from the precursor sequence by a
 autocatalytic mechanism.
 -1- SIMILARITY: Belongs to peptidase family C14.
 -1- SIMILARITY: Contains 1 CARD domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; L03799; AAA39306.1; -.
 EMBL; L28095; AAA20209.1; -.
 DR EMBL; U04269; AAA56992.1; -.
 DR EMBL; BC008152; AAH08152.1; -.
 DR PIR; A46495; A46495.
 DR HSP; P29466; 1ICE.

```
DR MEROPS; C14.001; -.
DR MGD; MGI:96544; Casp1.
DR InterPro; IPR001315; CARD.
DR InterPro; IPR002138; ICE_P10.
DR InterPro; IPR001309; ICE_P20.
DR InterPro; IPR002338; Peptidase_C14.
DR Pfam; PF00619; CARD; 1.
DR Pfam; PF00656; Peptidase_C14; 1.
DR PRINTS; PR00376; IL1RCENZYM.
DR SMART; SM00114; CARD; 1.
DR SMART; SM00115; CASG; 1.
DR PROSITE; PS02029; CARD; 1.
DR PROSITE; PS01122; CASPASE_CYS; 1.
DR PROSITE; PS01121; CASPASE_HIS; 1.
DR PROSITE; PS02027; CASPASE_P10; 1.
DR PROSITE; PS02028; CASPASE_P20; 1.
KW Hydrolase; Thiol protease; Zymogen.
FT PROPEP 1 ?118
FT CHAIN ?119 296
FT PROPEP 297 314
FT CHAIN 315 402
FT DOMAIN 1 91
FT ACT_SITE 236 236
FT ACT_SITE 284 284
FT CONFLICT 3 6
SQ SEQUENCE 402 AA; 45640 MW; 3BEBCEFA67C69B03 CRC64;

Query Match 65.4%; Score 34; DB 1; Length 402;
Best Local Similarity 100.0%; Pred.No. 44;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CPLEKA 7
DB 135 CPLEKA 140

RESULT 15
GSPF_ERWCH STANDARD; PRT; 404 AA.
AC P31704;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE General secretion pathway protein F (Pectic enzymes secretion protein
outF).
DE outF.
GN outF.
OS Erwinia chrysanthemi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Pectobacterium.
OX NCBI_TaxID=556;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EC16;
RX MEDLINE=93054355; PubMed=1429461;
RA Lindeberg M., Collmer A.;
RT "Analysis of eight out genes in a cluster required for pectic enzyme
secretion by Erwinia chrysanthemi: sequence comparison with secretion
genes from other Gram-negative bacteria."
RL J. Bacteriol. 174:7385-7397(1992).
CC -!- FUNCTION: INVOLVED IN A GENERAL SECRETION PATHWAY (GSP) FOR THE
CC EXPORT OF PROTEINS. REQUIRED FOR THE TRANSLOCATION OF THE
CC MULTIPLE PECTIC ENZYMES.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -!- SIMILARITY: BELONGS TO THE PULF/OUTF/EXEF/XPSF/XCPS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC EMBL; L02214; AAA24833.1; -.
DR PIR; D47021; D47021.
DR InterPro; IPR003004; Bac_GSPF.
DR InterPro; IPR001992; Bacf_secr_systII.
DR Pfam; PF00482; GSPF_P; 1.
DR PRINTS; PR00812; BCTERIALGSPF.
DR PROSITE; PS00874; T2SP_F; 1.
KW Transport; Transmembrane; Inner membrane.
FT TRANSMEM 172 192
FT TRANSMEM 223 243
FT TRANSMEM 375 395
SQ SEQUENCE 404 AA; 44751 MW; AB45B7D04B9B6FAE CRC64;

Query Match 65.4%; Score 34; DB 1; Length 404;
Best Local Similarity 77.8%; Pred.No. 44;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACPLKALD 9
DB 82 ALPLEEALD 90

Search completed: September 9, 2004, 10:53:49
Job time : 3.83871 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:44:32 ; Search time 5.09677 Seconds
(without alignments)
188.730 Million cell updates/sec

Title: US-10-067-618-4
Perfect score: 52
Sequence: 1 ACPLKALDV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78.*
1: Piri:**
2: Piri:**
3: Piri:**
4: Piri:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	101	2 A48219	calvasculin - huma
2	41	78.8	100	2 A53217	placental calcium-
3	38	73.1	340	1 JH0099	anthranilate phosph
4	38	73.1	401	2 S20035	pulF protein - Kle
5	36	69.2	101	2 S06207	calvasculin - mous
6	36	69.2	101	2 S01759	calvasculin - rat
7	36	69.2	126	2 AF1545	conserved hypothet
8	36	69.2	126	2 AG1187	conserved hypothet
9	36	69.2	231	2 T22869	hypothetical prote
10	36	69.2	766	2 D83404	5-methyltetrahydro
11	36	69.2	1051	2 T48933	WD repeat domain p
12	35	67.3	118	2 A83742	cinnamoyl ester hy
13	35	67.3	288	2 B89930	hypothetical prote
14	35	67.3	300	2 F90274	hypothetical prote
15	35	67.3	402	1 B71173	probable molybdopt
16	35	67.3	429	2 H98112	histidine-trNA lig
17	35	67.3	818	2 B83537	hypothetical prote
18	34	65.4	48	2 S32448	pol polyprotein -
19	34	65.4	87	2 B46200	retrovirus-related
20	34	65.4	87	2 S32444	pol polyprotein -
21	34	65.4	87	2 S32445	pol polyprotein -
22	34	65.4	87	2 S32447	pol polyprotein -
23	34	65.4	87	2 S32449	pol polyprotein -
24	34	65.4	128	2 B69779	conserved hypothet
25	34	65.4	141	1 QQVLM8	gene X protein - w
26	34	65.4	263	2 T27641	hypothetical prote
27	34	65.4	304	2 T24703	hypothetical prote
28	34	65.4	386	2 S41497	thyroid hormone re
29	34	65.4	388	2 JC7510	benzoate X recepto

ALIGNMENTS

RESULT 1

A48219
calvasculin - human
N:Alternate names: calcium-binding protein L (CAPL); metastasin; placental calcium-bindin
C:Species: Homo sapiens (man)
C>Date: 02-Jun-1994 #sequence revision 02-Jun-1994 #text_change 13-Aug-1999
C:Accession: A48219; B44470; F01223
R:Engelkamp, D.; Schaefer, B.W.; Mattei, M.G.; Erne, P.; Heizmann, C.W.
Proc. Natl. Acad. Sci. U.S.A. 90, 6547-6551, 1993
A:Title: Six S100 genes are clustered on human chromosome 1q21: identification of two ge
A:Reference number: A48219; MUID:93342029; PMID:8341667
A:Accession: A48219
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-101 <ENG>
A:Cross-references: GB:Z18950; NID:G396708; PIDN:CAA79474.1; PID:G396710
R:Engelkamp, D.; Schafer, B.W.; Erne, P.; Heizmann, C.W.
Biochemistry 31, 10258-10264, 1992
A:Title: S100 alpha, CAPL, and CACY: molecular cloning and expression analysis of three
A:Reference number: A44470; MUID:93041710; PMID:1384693
A:Accession: B44470
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-101 <EN2>
A:Cross-references: GB:M80563; GB:M77499; NID:G179916; PIDN:AAA51920.1; PID:G179917
A:Experimental source: heart
A:Note: sequence extracted from NCBI backbone (NCBIP:116496)
R:Tomida, Y.; Terasawa, M.; Kobayashi, R.; Hidaka, H.
Biochem. Biophys. Res. Commun. 189, 1310-1316, 1992
A:Title: Calcyclin and calvasculin exist in human platelets.
A:Reference number: PC1222; MUID:93129189; PMID:1482346
A:Accession: PC1223
A:Molecule type: protein
A:Residues: 8-18;36-57 <TOM>
A:Experimental source: platelets
C:Genetics:
A:Gene: GDB:S100A4; CAPL
A:Cross-references: GDB:119748; OMIM:114210
A:Map position: 1q21-1q21
A:Introns: 47/3
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; EF hand
F:7-41/Domain: calmodulin repeat homology <EF1>
F:50-82/Domain: calmodulin repeat homology <EF2>

Query Match 100.0% ; Score 52; DB 2; Length 101;
Best Local Similarity 100.0% ; Pred. No. 0.0055;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ACPLKALDV 10
DB 2 ACPLKALDV 11

Thu Sep 9 12:03:22 2004

```

RESULT 2
A53217
placental calcium-binding protein homolog - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Oct-1995 #sequence_revision 19-Oct-1995 #text_change 17-Mar-1999
C:Accession: A53217
R:Polans, A.S.; Palczewski, K.; Asson-Batres, M.A.; Ohguro, H.; Witkowska, D.; Haley, T.
J. Biol. Chem. 269, 6233-6240, 1994
A:Title: Purification and primary structure of cap1, an S-100-related calcium-binding pr
A:Reference number: A53217; MUID:94164991; PMID:8119967
A:Accession: A53217
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-100 <POL>
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: acetylated amino end; calcium binding; EF hand
F:6-40/Domain: calmodulin repeat homology <EF1>
F:1/Modified site: acetylated amino end (Ala) #status experimental

Query Match 78.8%; Score 41; DB 2; Length 100;
Best Local Similarity 90.0%; Pred. No. 0.81;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACPLEKALDV 10
| | | | |
Db 1 AYPLEKALDV 10
| | | | |

RESULT 3
JH0099
anthranilate phosphoribosyltransferase (EC 2.4.2.18) - Bacillus pumilus
C:Species: Bacillus pumilus
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: JH0099
R:Rivas, M.V.; Jarvis, E.D.; Rudner, R.
Gene 87, 71-78, 1990
A:Title: The structure of the trpE, trpD and 5' trpC genes of Bacillus pumilus.
A:Reference number: JH0099; MUID:90236301; PMID:2110100
A:Accession: JH0099
A:Molecule type: DNA
A:Residues: 1-340 <RIV>
A:Cross-references: GB:M36468; NID:gl374669; PIDN:AAB02273.1; PID:gl374670
C:Genetics:
A:Gene: trpD
C:Superfamily: anthranilate phosphoribosyltransferase; trpD homology
C:Keywords: glycosyltransferase; pentosyltransferase; tryptophan biosynthesis
F:4-332/Domain: trpD homology <TRD>

Query Match 73.1%; Score 38; DB 1; Length 340;
Best Local Similarity 70.0%; Pred. No. 11;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACPLEKALDV 10
| | | | |
Db 64 AAPMERALDV 73
| | | | |

RESULT 4
S20035
pufF protein - Klebsiella pneumoniae
C:Species: Klebsiella pneumoniae
C:Date: 19-Mar-1997 #sequence_revision 19-Jul-1997 #text_change 20-Aug-1999
C:Accession: S20035; S77951
R:Possoy, O.; d'Enfert, C.; Reyss, I.; Pugsley, A.P.
Mol. Microbiol. 6, 95-105, 1992
A:Title: Pullulanase secretion in Escherichia coli K-12 requires a cytoplasmic protein a
A:Reference number: S20034; MUID:92149318; PMID:1738317
A:Accession: S20035
A:Molecule type: DNA
A:Residues: 1-381 <POS>
A:Cross-references: EMBL:M32613

```

```

R:Pugsley, A.P.; D'Enfert, C.; Reyss, I.; Wandersman, C.
submitted to the EMBL Data Library, January 1991
A:Reference number: S77951
A:Accession: S77951
A:Molecule type: DNA
A:Residues: 1-140,161-401 <PUG>
A:Cross-references: EMBL:M32613; NID:gl49301; PIDN:AAA25128.1; PID:gl49305
C:Superfamily: secretion protein xcpS
C:Keywords: transmembrane protein
F:162-188/Domain: transmembrane #status predicted <TM1>
F:218-234/Domain: transmembrane #status predicted <TM2>
F:372-390/Domain: transmembrane #status predicted <TM3>

Query Match 73.1%; Score 38; DB 2; Length 401;
Best Local Similarity 88.9%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACPLEKALD 9
| | | | |
Db 77 AIPLEKALD 85
| | | | |

RESULT 5
S06207
calvasculin - mouse
N:Alternate names: calcium-binding protein mts1; calcium-binding protein pEL98; placenta
C:Species: Mus musculus (house mouse)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 20-Jun-2000
C:Accession: S06207; JH0097; S07981; A26803; A41411; I48674
R:Ebraldtze, A.; Tulchinsky, E.; Grigorian, M.; Afanasyeva, A.; Senin, V.; Revazova, E.;
Genes Dev. 3, 1086-1093, 1989
A:Title: Isolation and characterization of a gene specifically expressed in different me

A:Reference number: S06207; MUID:89378739; PMID:2550322
A:Accession: S06207
A:Molecule type: mRNA
A:Residues: 1-101 <EBR>
A:Cross-references: EMBL:X16190; NID:g54926; PIDN:CAA34316.1; PID:g54927
R:Tulchinsky, E.M.; Grigorian, M.S.; Ebraldtze, A.K.; Mlshina, N.I.; Lukanidin, E.M.
Gene 87, 219-223, 1990
A:Title: Structure of gene mts1, transcribed in metastatic mouse tumor cells.
A:Reference number: JH0097; MUID:90236313; PMID:2332170
A:Accession: JH0097
A:Molecule type: DNA
A:Residues: 1-101 <TUL>
A:Cross-references: GB:M36578; GB:M36579
A:Experimental source: liver
R:Tulchinsky, B. EMBL Data Library, August 1989
submitted to the EMBL Data Library, August 1989
A:Reference number: S07981
A:Accession: S07981
A:Molecule type: DNA
A:Residues: 1-47, 'VSGSXFNQ', 56-57, 'RTDEAA', <TUL>
A:Cross-references: EMBL:X16094; NID:g53249; PIDN:CAA34224.1; PID:g53250
R:Jackson-Grusby, L.B.; Swiergiel, J.; Linzer, D.I.H.
Nucleic Acids Res. 15, 6677-6690, 1987
A:Title: A growth-related mRNA in cultured mouse cells encodes a placental calcium bindi
A:Reference number: A26803; MUID:87316927; PMID:3628004
A:Accession: A26803
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-101 <JAC>
A:Cross-references: GB:X05835; NID:g50310; PIDN:CAA29282.1; PID:g50311
R:Goto, K.; Endo, H.; Fujiyoshi, T.
J. Biochem. 103, 48-53, 1988
A:Title: Cloning of the sequences expressed abundantly in established cell lines: ident
A:Reference number: A41411; MUID:88198109; PMID:3162911
A:Accession: A41411
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-101 <GOT>
A:Cross-references: GB:D00208; NID:g220569; PIDN:BAA00148.1; PID:g220570
R:Tulchinsky, E.; Kramerov, D.; Ford, H.L.; Reshetnyak, E.; Lukanidin, E.; Zain, S.

```

Oncogene 8, 79-86, 1993
A:Title: Characterization of a positive regulatory element in the mts1 gene.
A:Reference number: I48674; MUID:93141279; PMID:8423998
A:Accession: I48674
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-47, VSGSKFNGQ', 48-54 <RES>
A:Cross-references: EMBL:X16094; NID:G53249; PIDN:CAA34224.1; PID:G53250
C:Comment: Gene mts1 is expressed in metastatic cells.
C:Genetics:
A:Gene: mts1
A:Introns: 47/3
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; cancer; EF hand
F:7-41/Domain: calmodulin repeat homology <EF1>
F:50-82/Domain: calmodulin repeat homology <EF2>

Query Match 69.2%; Score 36; DB 2; Length 101;
Best Local Similarity 80.0%; Pred. No. 8;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACPLEKALDV 10
| | | | |
Db 2 ARPLEEALDV 11

RESULT 6
S01759
calvasculin - rat
N:Alternate names: calcium-binding protein p9Ka/42A; gene p9Ka protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 13-Aug-1999
C:Accession: S01759; B31373; S26496; S26497
R:Barracough, R.; Savin, J.; Dube, S.K.; Rudland, P.S.
J. Mol. Biol. 198, 13-20, 1987
A:Title: Molecular cloning and sequence of the gene for p9Ka : a cultured myoepithelial
A:Reference number: S01759; MUID:88118907; PMID:3430604
A:Accession: S01759
A:Molecule type: DNA
A:Residues: 1-101 <BAR>
A:Cross-references: EMBL:X06916; NID:G56832; PIDN:CAA30014.1; PID:G56833
R:Wasiakowski, P.; Shooter, E.M.
Proc. Natl. Acad. Sci. U.S.A. 85, 1277-1281, 1988
A:Title: Nerve growth factor induces the genes for two proteins related to a family of c
A:Reference number: A94189; MUID:88125019; PMID:3422491
A:Accession: B31373
A:Molecule type: mRNA
A:Residues: 1-101 <AS>
A:Cross-references: GB:J03628; NID:G206829; PIDN:AAA42098.1; PID:G206830
R:de Vouge, M.W.; Mukherjee, B.B.
Oncogene 7, 109-119, 1992
A:Title: Transformation of normal rat kidney cells by v-k-ras enhances expression of tra
A:Reference number: S26496; MUID:92158347; PMID:1741158
A:Accession: S26496
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 65-100 <DEV>
A:Cross-references: EMBL:X64023
A:Accession: S26497
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 3-97 <DE2>
A:Cross-references: EMBL:X64022
C:Genetics:
A:Gene: p9Ka
A:Introns: 47/3
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; EF hand
F:50-82/Domain: calmodulin repeat homology <EF2>

Query Match 69.2%; Score 36; DB 2; Length 101;
Best Local Similarity 80.0%; Pred. No. 8;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACPLEKALDV 10
| | | | |
Db 2 ARPLEEALDV 11

RESULT 6
S01759
calvasculin - rat
N:Alternate names: calcium-binding protein p9Ka/42A; gene p9Ka protein
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 13-Aug-1999
C:Accession: S01759; B31373; S26496; S26497
R:Barracough, R.; Savin, J.; Dube, S.K.; Rudland, P.S.
J. Mol. Biol. 198, 13-20, 1987
A:Title: Molecular cloning and sequence of the gene for p9Ka : a cultured myoepithelial
A:Reference number: S01759; MUID:88118907; PMID:3430604
A:Accession: S01759
A:Molecule type: DNA
A:Residues: 1-101 <BAR>
A:Cross-references: EMBL:X06916; NID:G56832; PIDN:CAA30014.1; PID:G56833
R:Wasiakowski, P.; Shooter, E.M.
Proc. Natl. Acad. Sci. U.S.A. 85, 1277-1281, 1988
A:Title: Nerve growth factor induces the genes for two proteins related to a family of c
A:Reference number: A94189; MUID:88125019; PMID:3422491
A:Accession: B31373
A:Molecule type: mRNA
A:Residues: 1-101 <AS>
A:Cross-references: GB:J03628; NID:G206829; PIDN:AAA42098.1; PID:G206830
R:de Vouge, M.W.; Mukherjee, B.B.
Oncogene 7, 109-119, 1992
A:Title: Transformation of normal rat kidney cells by v-k-ras enhances expression of tra
A:Reference number: S26496; MUID:92158347; PMID:1741158
A:Accession: S26496
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 65-100 <DEV>
A:Cross-references: EMBL:X64023
A:Accession: S26497
A>Status: preliminary; translation not shown
A:Molecule type: mRNA
A:Residues: 3-97 <DE2>
A:Cross-references: EMBL:X64022
C:Genetics:
A:Gene: p9Ka
A:Introns: 47/3
C:Superfamily: S-100 protein; calmodulin repeat homology
C:Keywords: calcium binding; EF hand
F:50-82/Domain: calmodulin repeat homology <EF2>

Qy 1 ACPLEKALDV 10
| | | | |
Db 2 ARPLEEALDV 11

RESULT 7
A1545
conserved hypothetical protein lin0902 [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AF1545
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H. Science 294, 849-852, 2001
A:Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AF1545
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-126 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96134.1; PID:G16413352; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin0902

Query Match 69.2%; Score 36; DB 2; Length 126;
Best Local Similarity 66.7%; Pred. No. 9.9;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACPLEKALD 9
| | | | |
Db 106 ACPVAKSLD 114

RESULT 8
AG1187
conserved hypothetical protein lmo0903 [imported] - Listeria monocytogenes (strain EGD-e)
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AG1187
R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Science 294, 849-852, 2001
A:Authors: Krefte, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1187
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-126 <GLA>
A:Cross-references: GB:NC_003210; PIDN:CAC98981.1; PID:G16410306; GSPDB:GN00177
A:Experimental source: strain EGD-e
C:Genetics:
A:Gene: lmo0903

Query Match 69.2%; Score 36; DB 2; Length 126;
Best Local Similarity 66.7%; Pred. No. 9.9;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACPLEKALD 9
| | | | |
Db 106 ACPVAKSLD 114

RESULT 9
T22869
hypothetical protein F57G4.9 - Caenorhabditis elegans

Thu Sep 9 12:03:22 2004

us-10-067-618-4.rpr

C;Species: *Caenorhabditis elegans*
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C;Accession: T22869
 R;Ainscough, R.
 submitted to the EMBL Data Library, November 1996
 A;Reference number: Z19629
 A;Accession: T22869
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-231 <WIL>
 A;Cross-references: EMBL:Z81554; PIDN: CAB04511.1; GSPDB:GN00023; CESP:F57G4.9
 A;Experimental source: clone F57G4
 C;Genetics:
 A;Gene: CESP:F57G4.9
 A;Map position: 5
 A;Introns: 76/1

Query Match 69.2%; Score 36; DB 2; Length 231;
 Best Local Similarity 55.6%; Pred. No. 18;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
 DB 177 CPYEKSVDI 185

RESULT 10
 DB3404
 5-methyltetrahydropteroyltriglutamate- homocysteine S-methyltransferase PA1927 [imported]
 C;Species: *Pseudomonas aeruginosa*
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C;Accession: D83404
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bradman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim, J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: D83404
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-766 <STO>
 A;Cross-references: GB:AE004619; GB:AE004091; NID:G9947920; PIDN:BA05315.1; GSPDB:GN001
 A;Experimental source: strain PA01
 C;Genetics:
 A;Gene: metE; PA1927
 C;Superfamily: cobalamin-independent methionine synthase

Query Match 69.2%; Score 36; DB 2; Length 766;
 Best Local Similarity 77.8%; Pred. No. 60;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
 DB 302 CDLEKALEV 310

RESULT 11
 T48933
 WD repeat domain protein - *Arabidopsis thaliana*
 N;Alternate names: protein F14L2.80
 C;Species: *Arabidopsis thaliana* (mouse-ear cress)
 C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C;Accession: T48933
 R;Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unseld, M.; Mewes, H.W.; Rudd, S.; submitted to the Protein Sequence Database, April 2000
 A;Reference number: Z25008
 A;Accession: T48933
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1051 <JOR>
 A;Cross-references: EMBL:AL353818; GSPDB:GN00061; ATSP:F14L2.80
 A;Experimental source: cultivar Columbia; BAC clone F14L2

C;Genetics:
 A;Gene: ATSP:F14L2.80
 A;Map position: 3
 A;Introns: 36/3; 121/3; 205/3; 336/2; 668/3; 739/1; 773/1; 853/2; 894/2; 932/3

Query Match 69.2%; Score 36; DB 2; Length 1051;
 Best Local Similarity 77.8%; Pred. No. 82;
 Matches 7; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACPLKALD 9
 DB 607 ACPLKALD 615

RESULT 12
 A83742
 cinnamoyl ester hydrolase BH0737 [imported] - *Bacillus halodurans* (strain C-125)
 C;Species: *Bacillus halodurans*
 C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
 C;Accession: A83742
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirata, N.
 Nucleic Acids Res. 28, 4317-4331, 2000
 A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
 A;Reference number: A83650; MUID:20512582; PMID:11058132
 A;Accession: A83742
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-118 <STO>
 A;Cross-references: GB:AP001509; GB:BA000004; NID:G10173176; PIDN:BA04456.1; GSPDB:GN00
 A;Experimental source: strain C-125
 C;Genetics:
 A;Gene: BH0737
 C;Superfamily: conserved hypothetical protein MTH1285

Query Match 67.3%; Score 35; DB 2; Length 118;
 Best Local Similarity 66.7%; Pred. No. 15;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
 DB 14 CPVEATLDV 22

RESULT 13
 B89930
 hypothetical protein SA1337 [imported] - *Staphylococcus aureus* (strain N315)
 C;Species: *Staphylococcus aureus*
 C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
 C;Accession: B89930
 R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogino, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Ma, A.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
 Lancet 357, 1225-1240, 2001
 A;Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
 A;Reference number: A89758; MUID:21311952; PMID:11418146
 A;Accession: B89930
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-288 <KUR>
 A;Cross-references: GB:BA000018; PID:G13701305; PIDN:BA042599.1; GSPDB:GN00149
 A;Experimental source: strain N315
 C;Genetics:
 A;Gene: SA1337

Query Match 67.3%; Score 35; DB 2; Length 288;
 Best Local Similarity 50.0%; Pred. No. 36;
 Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACPLKALDV 10
 DB 189 SCPLNGLEI 198

RESULT 14

P90274

hypothetical protein porB-1 [imported] - Sulfolobus solfataricus
 C:Species: Sulfolobus solfataricus
 C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
 C:Accession: F90274
 R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
 arrett, R.A.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.
 submitted to GenBank, April 2001
 A:Description: Sulfolobus solfataricus complete genome.
 A:Reference number: A93139
 A:Accession: F90274
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-300 <KUR>
 A:Cross-references: GB:AE006641; NID:gi13814401; PIDN:AAK41453.1; GSPDB:GN00155
 C:Genetics:
 A:Gene: porB-1
 C:Superfamily: pyruvate synthase beta chain

Query Match 67.3%; Score 35; DB 2; Length 300;
 Best Local Similarity 66.7%; Pred. No. 37;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
 ||: |||||
 Db 26 CPIPKELDV 34

RESULT 15

B71173

Probable molybdopterin biosynthesis moea protein - Pyrococcus horikoshii
 C:Species: Pyrococcus horikoshii
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
 C:Accession: B71173
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin-
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi
 DNA Res. 5, 55-76, 1998
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
 A:Reference number: A71000; MUID:98344137; PMID:9679194
 A:Accession: B71173
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-402 <KAW>
 A:Cross-references: GB:AF000002; NID:g3236129; PIDN:BAA29671.1; PID:g3256988
 A:Experimental source: strain OT3
 A>Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C:Genetics:
 A:Gene: PH0582
 C:Superfamily: molybdenum cofactor biosynthesis protein mceA-2

Query Match 67.3%; Score 35; DB 1; Length 402;
 Best Local Similarity 87.5%; Pred. No. 50;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 PLEKALDV 10
 |||||:
 Db 8 PLEKALEV 15

Search completed: September 9, 2004, 10:59:38
 Job time : 6.09677 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:53:18 ; Search time 13.4194 Seconds
(without alignments)
238.975 Million cell updates/sec

Title: US-10-067-618-4
Perfect score: 52
Sequence: 1 ACPLEXALDV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	10	13 US-10-067-618-4	Sequence 4, Appli
2	52	100.0	10	13 US-10-135-152-4	Sequence 4, Appli
3	52	100.0	101	9 US-09-393-433-1	Sequence 1, Appli
4	52	100.0	101	9 US-09-781-509-1	Sequence 1, Appli
5	52	100.0	101	12 US-10-087-192-1158	Sequence 1158, Ap
6	52	100.0	101	13 US-10-067-618-2	Sequence 2, Appli
7	52	100.0	101	13 US-10-135-152-2	Sequence 2, Appli
8	52	100.0	101	14 US-10-269-643-1	Sequence 2, Appli
9	43	82.7	854	12 US-10-205-331-9	Sequence 9, Appli
10	43	82.7	854	15 US-10-428-487-26	Sequence 26, Appli
11	39	75.0	166	12 US-10-424-599-274517	Sequence 274517,
12	38	73.1	461	15 US-10-767-701-43530	Sequence 43530, A
13	38	73.1	562	12 US-10-424-599-174551	Sequence 174551,
14	38	73.1	613	16 US-10-437-963-177070	Sequence 177070,
15	38	73.1	751	15 US-10-221-074-8	Sequence 8, Appli

16 37 71.2 106 12 US-10-424-599-212819 Sequence 212819,
17 36 69.2 101 9 US-09-393-433-2 Sequence 2, Appli
18 36 69.2 101 9 US-09-781-509-2 Sequence 2, Appli
19 36 69.2 101 14 US-10-269-643-2 Sequence 2, Appli
20 36 69.2 119 12 US-10-087-192-1155 Sequence 1155, Ap
21 35 67.3 53 14 US-10-029-386-34244 Sequence 34244, A
22 35 67.3 137 14 US-10-106-698-7693 Sequence 7693, Ap
23 35 67.3 187 16 US-10-437-963-175503 Sequence 175503,
24 35 67.3 209 15 US-10-094-749-1773 Sequence 1773, Ap
25 35 67.3 263 15 US-10-421-654-34 Sequence 34, Appli
26 35 67.3 429 9 US-09-815-242-13506 Sequence 13506, A
27 35 67.3 439 12 US-10-282-122A-74212 Sequence 74212, A
28 35 67.3 932 16 US-10-437-963-155092 Sequence 155092,
29 34 65.4 28 9 US-09-989-903-23 Sequence 23, Appli
30 34 65.4 28 14 US-10-068-564-23 Sequence 23, Appli
31 34 65.4 76 12 US-10-424-599-213736 Sequence 213736,
32 34 65.4 201 16 US-10-767-701-62495 Sequence 62495, A
33 34 65.4 259 16 US-10-437-963-146823 Sequence 146823,
34 34 65.4 279 12 US-10-243-552-843 Sequence 843, App
35 34 65.4 344 16 US-10-437-963-196435 Sequence 196435,
36 34 65.4 355 14 US-10-106-698-4313 Sequence 4313, Ap
37 34 65.4 386 14 US-10-153-827-2 Sequence 2, Appli
38 34 65.4 402 8 US-08-459-455-38 Sequence 38, Appli
39 34 65.4 402 9 US-09-888-243-14 Sequence 14, Appli
40 34 65.4 402 10 US-09-773-670-4 Sequence 4, Appli
41 34 65.4 402 14 US-10-123-529-4 Sequence 4, Appli
42 34 65.4 402 14 US-10-123-529-6 Sequence 6, Appli
43 34 65.4 438 12 US-10-330-472-2 Sequence 2, Appli
44 34 65.4 438 15 US-10-289-762-107 Sequence 107, App
45 34 65.4 575 12 US-10-403-571-80 Sequence 80, Appli

ALIGNMENTS

RESULT 1
US-10-067-618-4
; Sequence 4, Application US/10067618
; Publication No. US20020169110A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 78792YAI11-Z
; CURRENT APPLICATION NUMBER: US/10/067,618
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/298,625
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 10
; TYPE: PPT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:peptide
US-10-067-618-4

Query Match 100.0%; Score 52; DB 13; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.0037;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACPLEXALDV 10

Db 1 ACPLEXALDV 10

```

; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/09/781,509
; CURRENT FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 09/393,433
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-781-509-1

Query Match 100.0%; Score 52; DB 9; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACPLEKALDV 10
Db 2 ACPLEKALDV 11

RESULT 5
US-10-087-192-1158
; Sequence 1158, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 52945200122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1158
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-1158

Query Match 100.0%; Score 52; DB 12; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACPLEKALDV 10
Db 2 ACPLEKALDV 11

RESULT 6
US-10-067-618-2
; Sequence 2, Application US/10067618
; Publication No. US20020169110A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 78792YAI11-Z
; CURRENT APPLICATION NUMBER: US/10/067,618
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/298,625
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 08/468,942

```



```
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-067-618-2

Query Match      100.0%; Score 52; DB 13; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACPELKALDV 10
Db      2 ACPELKALDV 11

RESULT 7
US-10-135-152-2
; Sequence 1, Application US/10135152
; Publication No. US20020172680A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 7879ZYAI11-2
; CURRENT APPLICATION NUMBER: US/10/135,152
; PRIOR FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US/09/298,625
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-135-152-2

Query Match      100.0%; Score 52; DB 13; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACPELKALDV 10
Db      2 ACPELKALDV 11

RESULT 8
US-10-269-643-1
; Sequence 1, Application US/10269643
; Publication No. US20030100503A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/10/269,643
; CURRENT FILING DATE: 2002-10-11
```

```
; PRIOR APPLICATION NUMBER: US/09/781,509
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 09/393,433
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-269-643-1

Query Match      100.0%; Score 52; DB 14; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.042;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACPELKALDV 10
Db      2 ACPELKALDV 11

RESULT 9
US-10-205-331-9
; Sequence 9, Application US/10205331
; Publication No. US20040058326A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018199
; CURRENT APPLICATION NUMBER: US/10/205,331
; CURRENT FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 117
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Vacuolar assembly protein VSP41 gene
US-10-205-331-9

Query Match      82.7%; Score 43; DB 12; Length 854;
Best Local Similarity 77.8%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      2 CPLEKALDV 10
Db      638 CPLEKALEI 646

RESULT 10
US-10-428-487-26
; Sequence 26, Application US/10428487
; Publication No. US20040006780A1
; GENERAL INFORMATION:
; APPLICANT: RASTELLI, LUCA K.
; APPLICANT: GERBER, HANS-PETER
; TITLE OF INVENTION: VEGF-MODULATED GENES AND METHODS EMPLOYING THEM
; FILE REFERENCE: 09800080-0103
; CURRENT APPLICATION NUMBER: US/10/428,487
; CURRENT FILING DATE: 2003-05-02
; PRIOR APPLICATION NUMBER: 09/815,153
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,201
; PRIOR FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
```

; SEQ ID NO 26
; LENGTH: 854
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-428-487-26

Query Match 82.7%; Score 43; DB 15; Length 854;
Best Local Similarity 77.8%; Pred. No. 20;
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
Db 638 CPLEKALEI 646

RESULT 11

US-10-424-599-274517
; Sequence 274517, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 274517
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_89911C.1.pep
US-10-424-599-274517

Query Match 75.0%; Score 39; DB 12; Length 166;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACPLKALDV 10
Db 143 ACPLSRVLDI 152

RESULT 12

US-10-767-701-43530
; Sequence 43530, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 43530
; LENGTH: 461
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(461)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C46685_1.pep
US-10-767-701-43530

Query Match 73.1%; Score 38; DB 16; Length 461;
Best Local Similarity 66.7%; Pred. No. 93;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
Db 364 CPINKALDM 372

RESULT 13

US-10-424-599-174551
; Sequence 174551, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Cao Yongwei
; APPLICANT: Zhou Yihua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 174551
; LENGTH: 562
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_128638C.1.pep
US-10-424-599-174551

Query Match 73.1%; Score 38; DB 12; Length 562;
Best Local Similarity 77.8%; Pred. No. 1.1e+02;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
Db 167 CPLKALDV 175

RESULT 14

US-10-437-963-177070
; Sequence 177070, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 177070
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_74759C.1.pep
US-10-437-963-177070

Query Match 73.1%; Score 38; DB 16; Length 613;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10

```
Db      256 CPINKALDM 264
||: |||:
RESULT 15
US-10-221-074-8
; Sequence 8, Application US/10221074
; Publication No. US20030208791A1
; GENERAL INFORMATION:
; APPLICANT: Bryan, Gregory T.
; APPLICANT: Hwang, Der-Chyan
; APPLICANT: Maxwell, Carl A.
; APPLICANT: McGonigle, Brian
; APPLICANT: Potter, Susan M.
; TITLE OF INVENTION: ENZYMES INVOLVED IN TERITERPENE SYNTHESIS
; FILE REFERENCE: BB1438 PCT
; CURRENT APPLICATION NUMBER: US/10/221,074
; CURRENT FILING DATE: 2003-01-08
; PRIOR APPLICATION NUMBER: 60/188,054
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 8
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Triticum aestivum
US-10-221-074-8
Query Match      73.1%; Score 38; DB 15; Length 751;
Best Local Similarity 66.7%; Pred. No. 1.6e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY      2 CPLEKALDV 10
||: |||:
Db      363 CPINKALDM 371
Search completed: September 9, 2004, 11:04:45
Job time : 15.4194 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:47:43 ; Search time 5.6129 Seconds
(without alignments)
91.977 Million cell updates/sec

Title: US-10-067-618-4
Perfect score: 52
Sequence: 1 ACPLKALDV 10

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/2/iaa/6C COMB.pep.*
6: /cgn2_6/prodata/2/iaa/6D COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	100.0	10	1 US-08-190-560-4	Sequence 4, Appli
2	52	100.0	10	1 US-08-469-277-4	Sequence 4, Appli
3	52	100.0	10	2 US-08-468-946-4	Sequence 4, Appli
4	52	100.0	10	2 US-08-468-942-4	Sequence 4, Appli
5	52	100.0	10	4 US-09-298-625-4	Sequence 2, Appli
6	52	100.0	101	1 US-08-190-560-2	Sequence 2, Appli
7	52	100.0	101	1 US-08-469-277-2	Sequence 2, Appli
8	52	100.0	101	2 US-08-468-946-2	Sequence 2, Appli
9	52	100.0	101	2 US-08-468-942-2	Sequence 2, Appli
10	52	100.0	101	4 US-09-298-625-2	Sequence 2, Appli
11	36	69.2	301	4 US-09-134-001C-3220	Sequence 3220, Ap
12	36	69.2	794	4 US-09-252-991A-31824	Sequence 31824, A
13	35	67.3	369	1 US-08-844-055-2	Sequence 2, Appli
14	35	67.3	369	3 US-09-006-849-2	Sequence 2, Appli
15	35	67.3	417	4 US-09-489-039A-12272	Sequence 2, Appli
16	35	67.3	429	1 US-08-906-744A-2	Sequence 2, Appli
17	35	67.3	429	3 US-09-093-134-2	Sequence 2, Appli
18	35	67.3	822	4 US-09-252-991A-29534	Sequence 29534, A
19	34	65.4	28	4 US-09-187-789-23	Sequence 23, Appli
20	34	65.4	28	4 US-09-139-600-18	Sequence 23, Appli
21	34	65.4	177	4 US-09-489-039A-7497	Sequence 18, Appli
22	34	65.4	287	1 US-08-365-981-10	Sequence 10, Appli
23	34	65.4	287	1 US-08-365-981-11	Sequence 10, Appli
24	34	65.4	386	4 US-08-875-082-2	Sequence 11, Appli
25	34	65.4	402	2 US-08-394-189B-14	Sequence 2, Appli
26	34	65.4	402	3 US-08-258-287B-38	Sequence 14, Appli
27	34	65.4	402	3 US-08-368-704C-38	Sequence 38, Appli

28 34 65.4 402 3 US-08-954-536-16 Sequence 16, Appli
29 34 65.4 402 3 US-08-908-436-4 Sequence 4, Appli
30 34 65.4 402 4 US-09-291-289-6 Sequence 6, Appli
31 34 65.4 402 5 PCT-US93-05705-14 Sequence 14, Appli
32 34 65.4 438 4 US-09-198-452A-107 Sequence 107, Appli
33 34 65.4 587 2 US-08-871-266B-18 Sequence 18, Appli
34 34 65.4 587 2 US-09-018-864A-18 Sequence 18, Appli
35 34 65.4 587 3 US-08-871-267B-24 Sequence 24, Appli
36 34 65.4 587 3 US-09-618-419-24 Sequence 24, Appli
37 34 65.4 804 4 US-09-107-532A-6348 Sequence 6348, Ap
38 33 63.5 89 1 US-07-987-272A-10 Sequence 10, Appli
39 33 63.5 90 4 US-09-370-838-200 Sequence 200, Appli
40 33 63.5 92 4 US-09-621-976-7537 Sequence 7537, Ap
41 33 63.5 157 4 US-09-107-532A-4454 Sequence 4454, Ap
42 32 61.5 157 4 US-09-252-991A-25900 Sequence 25900, A
43 32 61.5 230 4 US-09-328-352-6355 Sequence 6355, Ap
44 32 61.5 348 1 US-08-459-489-10 Sequence 10, Appli
45 32 61.5 348 1 US-08-458-686-10 Sequence 10, Appli

ALIGNMENTS

RESULT 1
US-08-190-560-4
; Sequence 4, Application US/08190560
; Patent No. 5798257
; GENERAL INFORMATION:
; APPLICANT: Zain, Saveeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-190-560-4

Query Match 100.0%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ACPLKALDV 10
Db 1 ACPLKALDV 10

ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,946
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,560
FILING DATE: 31-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 78792Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-468-946-4

Query Match	100.0%;	Score 52;	DB 2;	Length 10;
Best Local Similarity	100.0%;	Pred. No. 0.00096;		
Mismatches	0;	Mismatches	0;	Indels
Conservative	10;			

Qy 1 ACPLEKALDV 10
|||
nb 1 ACPLEKALDV 10

RESULT 4
US-08-468-942-4
; Sequence 4, Application US/08468942
; Patent No. 5965360
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza

CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,942
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,560
FILING DATE: 31-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Digigallo, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 78792Y
TELECOMMUNICATION INFORMATION:

```

RESULT 2
US-08-469-277-4
; Sequence 4, Application US/08469277
; Patent No. 5801142
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE WTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,277
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 78792Y
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE:
; US-08-469-277-4

```

```
Query Match          100.0%; Score 52; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 10: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 1 ACPLEKALDV 10
| | | | | | | |
Db 1 ACPLEKALDV 10

```

RESULT 3
US-08-468-946--4
; Sequence 4, Application US/08468946
; Patent No. 5843686
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States

```

TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-468-942-4

Query Match 100.0%; Score 52; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACPLKALDV 10
|||||
Db 1 ACPLKALDV 10

RESULT 5

US-09-298-625-4
Sequence 4, Application US/09298625
Patent No. 6638504
GENERAL INFORMATION:
APPLICANT: Lukanidin, Eugene
TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
CURRENT APPLICATION NUMBER: US/09/298,625
CURRENT FILING DATE: 1999-04-23
PRIOR APPLICATION NUMBER: 08/468,942
PRIOR FILING DATE: 1995-06-06
PRIOR APPLICATION NUMBER: 08/190,560
PRIOR FILING DATE: 1994-01-31
PRIOR APPLICATION NUMBER: 07/981,455
PRIOR FILING DATE: 1992-11-25
PRIOR APPLICATION NUMBER: 07/550,600
PRIOR FILING DATE: 1990-07-09
NUMBER OF SEQ ID NOS: 17
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 4
LENGTH: 10
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: peptide
US-09-298-625-4

Query Match 100.0%; Score 52; DB 4; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACPLKALDV 10
|||||
Db 1 ACPLKALDV 10

RESULT 6

US-08-190-560-2
Sequence 2, Application US/08190560
Patent No. 5798257
GENERAL INFORMATION:
APPLICANT: Zain, Sayeeda
APPLICANT: Lukanidin, Eugene
TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
THE MTS-1 GENE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York

COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/190,560
FILING DATE:
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 78792Y
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 101 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-190-560-2

Query Match 100.0%; Score 52; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACPLKALDV 10
|||||
Db 2 ACPLKALDV 11

RESULT 7

US-08-469-277-2
Sequence 2, Application US/08469277
Patent No. 5801142
GENERAL INFORMATION:
APPLICANT: Zain, Sayeeda
APPLICANT: Lukanidin, Eugene
TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
THE MTS-1 GENE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSER: SCULLY, SCOTT, MURPHY & PRESSER
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,277
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/190,560
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 78792Y
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366

```

; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-277-2

Query Match      100.0%; Score 52; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACPLEKALDV 10
Db      2 ACPLEKALDV 11
      |||||
RESULT 8
US-08-468-946-2
; Sequence 2, Application US/08468946
; Patent No. 5843686
; GENERAL INFORMATION:
; APPLICANT: Zain, Saveeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/468,946
; FILING DATE: 06-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/190,560
; FILING DATE: 31-JAN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: DiGiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 7879ZY
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-468-946-2

Query Match      100.0%; Score 52; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 ACPLEKALDV 10
Db      2 ACPLEKALDV 11
      |||||

```


; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-298-625-2

Query Match 100.0%; Score 52; DB 4; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ACPEKALDV 10
Db 2 ACPEKALDV 11

RESULT 11
US-09-134-001C-3220
; Sequence 3220, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007

; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3220
; LENGTH: 301
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3220

Query Match 69.2%; Score 36; DB 4; Length 301;
Best Local Similarity 50.0%; Pred. No. 34;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ACPEKALDV 10
Db 197 SCPEQGLEI 206

RESULT 12
US-09-252-991A-31824
; Sequence 31824, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31824
; LENGTH: 794
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31824

Query Match 69.2%; Score 36; DB 4; Length 794;
Best Local Similarity 77.8%; Pred. No. 94;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 2 CPLEKALDV 10
Db 330 CDLEKALEV 338

RESULT 13
US-08-844-055-2
; Sequence 2, Application US/08844055
; Patent No. 5747313
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 5747313el Compounds
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,055
; FILING DATE: 18-APR-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 9607993.4
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimmi, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31457-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-844-055-2

Query Match 67.3%; Score 35; DB 1; Length 369;
Best Local Similarity 70.0%; Pred. No. 66;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACPEKALDV 10
Db 322 ALPIENALDV 331

RESULT 14
US-09-006-849-2
; Sequence 2, Application US/09006849
; Patent No. 6071731
; GENERAL INFORMATION:
; APPLICANT: Lawlor, Elizabeth
; TITLE OF INVENTION: No. 6071731el Compounds
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA

;
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,849
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/844,055
; FILING DATE: 18-APR-1997
; APPLICATION NUMBER: 9607993.4
; FILING DATE: 18-APR-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Gimm1, Edward R
; REGISTRATION NUMBER: 38,891
; REFERENCE/DOCKET NUMBER: P31457-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-4478
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 369 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-006-849-2

Query Match 67.3%; Score 35; DB 3; Length 369;
Best Local Similarity 70.0%; Pred. No. 66;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 ACPLKALDV 10
| : | | | |
Db 322 ALPIENALDV 331

RESULT 15
US-09-489-039A-12272
; Sequence 12272, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12272
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-12272

Query Match 67.3%; Score 35; DB 4; Length 417;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PLEKALD 9
| : | | | |
Db 95 PLEKALD 101

Search completed: September 9, 2004, 11:01:10
Job time : 5.6129 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:32:01 ; Search time 21.0323 Seconds
(without alignments)
134.340 Million cell updates/sec

Title: US-10-067-618-4

Perfect score: 52

Sequence: 1 ACPLKALDV 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Geneseq_29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match Length	ID	Description
1	52	100.0	10	2 AAR80454	Aar80454 Human mts
2	52	100.0	10	3 AAB37433	Aab37433 Human mts
3	52	100.0	10	6 ABU08518	Abu08518 Human mts
4	52	100.0	101	2 AAR20560	Aar20560 Human mts
5	52	100.0	101	2 AAR80453	Aar80453 Human mts
6	52	100.0	101	3 AAB45534	Aab45534 Human S10
7	52	100.0	101	3 AAB37432	Aab37432 Human mts
8	52	100.0	101	4 AAB72386	Aab72386 Human mts
9	52	100.0	101	6 ABU08513	Abu08513 Human mts
10	52	100.0	101	7 ADD14157	Add14157 Human src
11	43	82.7	854	6 ABM04788	Abm04788 Human vac
12	43	82.7	854	6 ABB62781	Abb62781 Human vas
13	39	75.0	539	4 ABB59576	Abb59576 Drosophil
14	38	73.1	751	4 AAB09797	Aae09797 Wheat oxy
15	36	69.2	101	4 AAB72387	Aab72387 Murine Mt
16	36	69.2	126	5 ABB49477	Abb49477 Listeria
17	36	69.2	296	4 AAG82290	Aag82290 S. epider
18	36	69.2	301	5 ABB38375	Abb38375 Staphyloc
19	36	69.2	418	4 ABB71839	Abb71839 Drosophil
20	35	67.3	115	5 ABB04124	Abb04124 Human ORF
21	35	67.3	137	4 AAG76919	Aag76919 Human col
22	35	67.3	209	6 ADA54205	Ada54205 Human pro
23	35	67.3	297	6 ABBM71065	Abm71065 Staphyloc
24	35	67.3	347	6 AAC26259	Aac26259 MDDT rela
25	35	67.3	369	2 AAW31903	Aaw31903 Streptoco

ALIGNMENTS

RESULT 1
AAR80454

ID AAR80454 standard; peptide; 10 AA.

XX AC AAR80454;

XX XX

DT 27-DEC-1995 (first entry)

DE Human mts-1 (2-11) peptide 1.

XX Metastasis; cancer; mts-1 gene; tumour; therapy; antigen; antibody.

XX OS Synthetic.

XX WO9520656-Al.

XX 03-AUG-1995.

XX 31-JAN-1995; 95WO-US001214.

XX 31-JAN-1994; 94US-00190560.

XX (RESE) RESEARCH CORP TECHNOLOGIES INC.

XX Zain S, Lukanidin E;

XX WPI; 1995-275441/36.

XX Nucleic acid encoding human mts-1, antigenic fragments and antibodies - useful for diagnosis of malignant cancer and metastatic potential of tumour cells.

XX Disclosure; Page 24; 124pp; English.

XX Antigenic peptides 1-4 (given in AAR80454-57) of human mts-1 protein (AAR80453) were used to raise polyclonal and monoclonal antibodies that detect the presence of mts-1 in tissue samples, esp. metastatic cells. Peptides 1, 3 and 4 are unique to mts-1, while peptide 2 generates antibodies reactive with many calcium binding proteins

XX Sequence 10 AA;

Query Match 100.0%; Score 52; DB 2; Length 10;

Best Local Similarity 100.0%; Pred. No. 0.0027;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACPLKALDV 10

|||||||

Aaw70983 Histidyl
Aay85136 Histidyl
Aau37913 Streptoco
Aau46288 Protein e
Abr59616 S. pneumo
Ade72439 Human end
Abb67659 Drosophil
Abg044851 Novel hum
Aam00899 Human bon
Abo00825 Polypepti
Ade09446 Novel pro
Aau30290 Novel hum
Aag73539 Human col
Aay98521 Xenopus o
Aay05379 Human HCM
Aaw94090 Wild-type
Aaw94091 Murine mu
Aab43233 Human ORF
Aab14249 Mouse ICE
Aay36878 Chlamydia

26 35 67.3 429 2 AAW70983
27 35 67.3 429 3 AAY85136
28 35 67.3 429 4 AAU37913
29 35 67.3 429 6 ABU46288
30 35 67.3 429 6 ABR59616
31 35 67.3 946 7 ADE72439
32 35 67.3 1340 4 ABB67659
33 34 65.4 160 4 ABG044851
34 34 65.4 279 4 AAM00899
35 34 65.4 279 6 ABO00825
36 34 65.4 279 7 ADE09446
37 34 65.4 283 4 AAU30290
38 34 65.4 355 4 AAG73539
39 34 65.4 386 2 AAY98521
40 34 65.4 402 2 AAY05379
41 34 65.4 402 2 AAW94090
42 34 65.4 402 3 AAW94091
43 34 65.4 402 3 AAB43233
44 34 65.4 402 3 AAB14249
45 34 65.4 438 2 AAY36878

Thu Sep 9 12:03:21 2004

us-10-067-618-4.rag

```

Db      1 ACPLEKALDV 10
RESULT 2
AAB37433
ID      AAB37433 standard; peptide; 10 AA.
XX
XX      AAB37433;
XX
XX      21-FEB-2001 (first entry)
DE
DE      Human mts-1 peptide fragment #1.
XX
XX      Human; mts-1; cytostatic; cancer; metastasis.
XX
XX      Homo sapiens.
XX      WO200064475-A1.
XX
XX      02-NOV-2000.
XX
XX      20-APR-2000; 2000WO-US011006.
XX
XX      23-APR-1999; 99US-00298625.
XX
XX      (RESE ) RESEARCH CORP TECHNOLOGIES INC.
XX      Lukanidin E;
XX      WPI; 2000-687266/67.
XX
XX      Treating cancer and preventing metastasis comprises administration of an
XX      anti-mts-1 protein antibody or antisense oligonucleotide.
XX
XX      Example 9; Page 27; 155pp; English.
XX
XX      The present invention relates to methods for treating cancer and
XX      preventing metastases, comprising the administration of a composition
XX      directed against the mts-1 protein (see AAC68131-C68132 and AAB37432).
XX      Mts-1 protein is a calcium-binding protein, and is thought to have a role
XX      in myoepithelial cell differentiation. The present sequence is a peptide
XX      fragment of the human mts-1 protein. This peptide was used to generate
XX      antibodies against mts-1 protein, which can be used to detect mts-1
XX      protein in clinical specimens
XX
XX      Sequence 10 AA;
XX
XX      Query Match      100.0%; Score 52; DB 3; Length 10;
XX      Best Local Similarity 100.0%; Pred. No. 0.0027;
XX      Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      1 ACPLEKALDV 10
XX      |||||
XX      Db      1 ACPLEKALDV 10
XX
XX      RESULT 3
XX      ID      ABU08518 standard; peptide; 10 AA.
XX
XX      AC      ABU08518;
XX
XX      27-MAY-2003 (first entry)
XX
XX      Human mts-1 protein, antigenic peptide #1.
XX
XX      Human; mts-1; metastatic cancer; p53; cytostatic; gene therapy; antigen.
XX
XX      Homo sapiens.
XX
XX      US2002172680-A1.
XX
XX      21-NOV-2002.
XX
XX      PD
XX
XX      29-APR-2002; 2002US-00135152.
XX
XX      09-JUL-1990; 90US-00550600.
XX
XX      25-NOV-1992; 92US-00981455.
XX
XX      31-JAN-1994; 94US-00190560.
XX
XX      06-JUN-1995; 95US-00468942.
XX
XX      23-APR-1999; 99US-00298635.
XX
XX      (LUKA/) LUKANIDIN E.
XX
XX      Lukanidin E;
XX
XX      WPI; 2003-328422/31.
XX
XX      Treating cancer by administering a reagent directed against the mts-1
XX      protein or an oligonucleotide capable of binding to mts-1 mRNA.
XX
XX      Example 9; Page 19; 64pp; English.
XX
XX      The invention relates to treating cancer comprising administering a
XX      reagent directed against the mts-1 (not defined) protein or an
XX      oligonucleotide capable of binding to mts-1 mRNA. Also include are a
XX      method of inactivating, destroying or nullifying a mts-1 protein or cells
XX      and a method of inhibiting metastasis in a cancerous cell. Experiments
XX      showed that mts-1 inhibited the phosphorylation of full-size p53 and the
XX      C-terminal protein fragment by PKC. Addition of the same concentrations
XX      of mts-1 to the PKC reaction mixture did not affect the phosphorylation
XX      of the N-terminal and DNA-binding domains of p53. The method is useful
XX      for treating cancers associated with the mts-1 gene. The present sequence
XX      represents a human mts-1 protein antigenic peptide
XX
XX      Sequence 10 AA;
XX
XX      Query Match      100.0%; Score 52; DB 6; Length 10;
XX      Best Local Similarity 100.0%; Pred. No. 0.0027;
XX      Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX      QY      1 ACPLEKALDV 10
XX      |||||
XX      Db      1 ACPLEKALDV 10
XX
XX      RESULT 4
XX      ID      AAR20560 standard; protein; 101 AA.
XX
XX      AC      AAR20560;
XX
XX      25-MAR-2003 (revised)
XX      14-MAY-1992 (first entry)
XX
XX      Human mts protein.
XX
XX      Metastatic cancer; antibodies; mouse; lung; liver; kidney; thyroid;
XX      breast cancer; cell growth.
XX
XX      Homo sapiens.
XX
XX      Key      Location/Qualifiers
XX      FT      Protein      1..101
XX      /label= mts-1
XX      FT      Peptide      2..11
XX      /note= "antigenic"
XX      FT      Peptide      22..37
XX      /note= "antigenic; calcium binding domain"
XX      FT      Peptide      42..54
XX      /note= "antigenic"
XX      FT      Peptide      87..101
XX      /note= "antigenic"
XX
XX      WO9200757-A.
XX
XX

```

PD 23-JAN-1992.
 XX
 PF 09-JUL-1990; 90US-00550600.
 XX
 PR 09-JUL-1990; 90US-00550600.
 XX
 PA (RESE) RESEARCH CORP TECHNOLOGIES INC.
 XX
 PI Zain S, Lukanidin E;
 XX
 DR WPI; 1992-056647/07.
 DR N-PSDB; AAQ20506.
 XX
 PT Metastatic cancer diagnosis by detection of mts-1 gene or protein - using
 PT antibody treatment of cancer and tumours of e.g. kidney, thyroid, lung
 PT and liver.
 XX
 PS Claim 9; Fig 2; 82pp; English.
 XX
 CC The sequence was deduced from the DNA sequence obtd. by screening a human
 CC cDNA library with mouse mts-1 cDNA probes. The antigenic Mts-1 peptides
 CC (see features) derived from the protein and anti- bodies raised to them
 CC are useful in the diagnosis of metastatic cancer, e.g. lung, kidney,
 CC thyroid or breast cancer. The peptide comprising the calcium binding site
 CC generates antibodies reactive with many members of the calcium binding
 CC protein family; the other three peptides are unique to mts-1 and generate
 CC antibodies specific only for this protein. Cell lines capable of
 CC expressing mts-1 are useful as model systems for in vitro and in vivo
 CC anti-metastasis drug screening. Pharmaceutical compns. contg. the mts-1
 CC protein or anti-cancer reagents may be used to promote cell growth, or
 CC for treating cancer, respectively. (Updated on 25-MAR-2003 to correct PA
 CC field.)
 XX
 SQ Sequence 101 AA;
 XX
 QY Query Match 100.0%; Score 52; DB 2; Length 101;
 Db Best Local Similarity 100.0%; Pred. No. 0.031;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACPEKALDV 10
 Db |||||
 2 ACPEKALDV 11
 RESULT 5
 AAR80453
 ID AAR80453 standard; protein; 101 AA.
 XX
 AC AAR80453;
 XX
 DT 27-DEC-1995 (first entry)
 XX
 DE Human mts-1 protein.
 XX
 KW Metastasis; cancer; mts-1 gene; tumour; therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO9520656-A1.
 XX
 PD 03-AUG-1995.
 XX
 PF 31-JAN-1995; 95WO-US001214.
 XX
 PR 31-JAN-1994; 94US-00190560.
 XX
 PA (RESE) RESEARCH CORP TECHNOLOGIES INC.
 XX
 PI Zain S, Lukanidin E;
 XX
 DR WPI; 1995-275441/36.
 DR N-PSDB; AAQ99177.
 XX

PT Nucleic acid encoding human mts-1, antigenic fragments and antibodies -
 PT useful for diagnosis of malignant cancer and metastatic potential of
 PT tumour cells.
 XX
 PS Claim 12; Page 92; 124pp; English.
 XX
 CC A human cDNA library was constructed in lambda-gt10 using poly(A)+ RNA
 CC prep'd. from Hela cells. The library was screened with a 32P- labeled
 CC mouse mts-1 vDNA probe. A clone was obtd. which comprised the full-length
 CC human mts-1 gene. The encoded protein is used for the diagnosis or
 CC therapy of cancer, and to raise antibodies
 XX
 SQ Sequence 101 AA;
 XX
 QY Query Match 100.0%; Score 52; DB 2; Length 101;
 Db Best Local Similarity 100.0%; Pred. No. 0.031;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACPEKALDV 10
 Db |||||
 2 ACPEKALDV 11
 RESULT 6
 AAB45534
 ID AAB45534 standard; protein; 101 AA.
 XX
 AC AAB45534;
 XX
 DT 22-FEB-2001 (first entry)
 XX
 DE Human S100A4 protein.
 XX
 KW S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;
 KW calcium-binding protein; calcium homeostasis; cardiac muscle;
 KW pumping capacity; myocardial cell; systolic calcium ion release;
 KW sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;
 KW valve defect.
 XX
 OS Homo sapiens.
 XX
 PN DE19915485-A1.
 XX
 PD 19-OCT-2000.
 XX
 PF 07-APR-1999; 99DE-01015485.
 XX
 PR 07-APR-1999; 99DE-01015485.
 XX
 PA (KATU/) KATUS H A.
 PA (REMP/) REMPPIS A.
 XX
 PI Katus HA, Remppis A;
 XX
 DR WPI; 2000-673510/66.
 DR N-PSDB; AAC81804.
 XX
 PT Composition containing S100 protein, corresponding nucleic acid or
 PT vector, useful for treating cardiomyopathy and cardiac insufficiency.
 XX
 PS Claim 35; Page 11; 36pp; German.
 XX
 CC This invention describes a novel composition for treating primary or
 CC secondary cardiomyopathy or cardiac insufficiency contains at least one
 CC S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or
 CC fragments, or a gene transfer vector containing (II), optionally
 CC formulated with auxiliaries and/or carriers. (I) are calcium-binding
 CC proteins involved in calcium homeostasis, so their overexpression in
 CC cardiac muscle will improve pumping capacity (and overall capacity) of
 CC the heart. In cultured myocardial cells they increase the contraction and
 CC relaxation rates associated with increased systolic calcium ion release
 CC from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are
 CC used to treat cardiomyopathy (CMP) where inherited or caused by

CC spontaneous mutations and ischemic CMP caused by arteriosclerosis.
 CC dilative CMP caused by toxic/infectious disease, cardiac disease caused
 CC by pulmonary and/or arterial hypertension, and structural disease caused
 CC by rhythm disorders or valve defects, generally any condition associated
 CC with reduced contractile force. Unlike calmodulin, which is expressed
 CC ubiquitously, (1) show tissue-specific expression and treat the
 CC underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac
 CC disease
 CC XX
 SQ Sequence 101 AA;

Query Match 100.0%; Score 52; DB 3; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.031;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACPLKALDV 10
 |||||
 Db 2 ACPLKALDV 11

RESULT 7
 AAB37432
 ID AAB37432 standard; protein; 101 AA.

XX AAB37432;
 XX 21-FEB-2001 (first entry)
 DT Human mts-1.
 DE Human; mts-1; cytostatic; cancer; metastasis.
 XX Homo sapiens.

OS
 PN WO200064475-A1.
 XX 02-NOV-2000.

XX 20-APR-2000; 2000WO-US011006.
 PF 23-APR-1999; 99US-00298625.

XX (RESE) RESEARCH CORP TECHNOLOGIES INC.
 PA Lukanidin E;

PI WPI; 2000-687266/67.
 XX N-PSDB; AAC68131, AAC68132.

DR Treating cancer and preventing metastasis comprises administration of an
 PT anti-mts-1 protein antibody or antisense oligonucleotide.

XX Example 6; Fig 2; 155pp; English.

XX The present sequence is human mts-1. Mts-1 protein is a calcium-binding
 CC protein, and is thought to have a role in myoepithelial cell
 CC differentiation. The present invention relates to methods for treating
 CC cancer and preventing metastases, comprising the administration of a
 CC composition directed against the mts-1 protein

XX Sequence 101 AA;

Query Match 100.0%; Score 52; DB 3; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.031;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACPLKALDV 10
 |||||
 Db 2 ACPLKALDV 11

RESULT 8
 AAB72386

ID AAB72386 standard; protein; 101 AA.
 XX AAB72386;

XX 24-MAY-2001 (first entry)

DT Human Mts 1 protein amino acid sequence.

DE Mts 1; Mts1-del75; Mts1-4S; S100A4; neuronal cell growth; cardiac arrest;
 XX neuronal degeneration; Parkinson's disease; Alzheimer's disease;
 KW Down's syndrome; stroke; sciatic crush; spinal cord injury; human.

XX Homo sapiens.

XX WO200118043-A2.

XX 15-MAR-2001.

XX 07-SEP-2000; 2000WO-US024495.

XX 10-SEP-1999; 99US-00393433.

XX (PROL-) PROLIFIA INC.

XX Bock E, Lukanidin EM, Berezin V;

XX WPI; 2001-235188/24.

XX New isolated functional derivatives of Mts protein for stimulating
 PT neuronal cell growth to treat neurological conditions e.g. Parkinson's
 PT disease, Alzheimer's disease, Down's syndrome and stroke.

XX Disclosure; Page; 60pp; English.

XX This invention relates to functional derivatives of an Mts protein,
 CC particularly Mts1-del75, and Mts1-4S. Mts 1 (also known as S100A4) is a
 CC member of the S100 gene family. Mts1 proteins and their functional
 CC derivatives are used to stimulate the growth of neuronal cells and in the
 CC treatment of neurological conditions including those characterised by
 CC neuronal degeneration, death or injury such as Parkinson's disease,
 CC Alzheimer's disease, Down's syndrome, stroke, cardiac arrest, sciatic
 CC crush, spinal cord injury, injury to sensory neurons or degenerative
 CC disease of the retina. The present sequence represents the human Mts 1
 CC protein, from which the Mts proteins of the invention may be derived.
 CC NOTE: The present sequence is not shown in the specification but it has
 CC been derived from the sequence given in US patent 5801142 (referred to on
 CC page 10 of the specification)

XX Sequence 101 AA;

Query Match 100.0%; Score 52; DB 4; Length 101;
 Best Local Similarity 100.0%; Pred. No. 0.031;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACPLKALDV 10
 |||||
 Db 2 ACPLKALDV 11

RESULT 9

ABU08513
 ID ABU08513 standard; protein; 101 AA.

XX AC ABU08513;

XX 27-MAY-2003 (first entry)

XX Human mts-1 protein.

XX Human; mts-1; metastatic cancer; p53; cytostatic; gene therapy.

XX Homo sapiens.

XX

PN US2002172680-A1.
XX
PD 21-NOV-2002.
XX
XX
XX 29-APR-2002; 2002US-00135152.
XX
PR 09-JUL-1990; 90US-00550600.
PR 28-NOV-1992; 92US-00981455.
PR 31-JAN-1994; 94US-00190560.
PR 06-JUN-1995; 95US-00468942.
PR 23-APR-1999; 99US-00298635.
XX
XX (LUKA/) LUKANIDIN E.
XX
XX Lukanidin E;
XX
XX WPI; 2003-328422/31.
DR N-PSDB; ABX93578.
XX
XX Treating cancer by administering a reagent directed against the mts-1
PT protein or an oligonucleotide capable of binding to mts-1 mRNA.
XX
XX Example 6; Fig 2; 64pp; English.
XX
XX The invention relates to treating cancer comprising administering a
CC reagent directed against the mts-1 (not defined) protein or an
CC oligonucleotide capable of binding to mts-1 mRNA. Also include are a
CC method of inactivating, destroying or nullifying a mts-1 protein or cells
CC and a method of inhibiting metastasis in a cancerous cell. Experiments
CC showed that mts-1 inhibited the phosphorylation of full-size p53 and the
CC C-terminal protein fragment by PKC. Addition of the same concentrations
CC of mts-1 to the PKC reaction mixture did not affect the phosphorylation
CC of the N-terminal and DNA-binding domains of p53. The method is useful
CC for treating cancers associated with the mts-1 gene. The present sequence
CC represents human mts-1 protein
XX
SQ Sequence 101 AA;

Query Match 100.0%; Score 52; DB 6; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACPLEKALDV 10
Db |||||
2 ACPLEKALDV 11

RESULT 10
ADD14157
ID ADD14157 standard; protein; 101 AA.
XX
AC ADD14157;
XX
DT 01-JAN-2004 (first entry)
XX
XX Human src biomarker polypeptide SEQ ID NO.346.
DE
XX predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human.
XX
XX Homo sapiens.
XX
XX WO2003062395-A2.
PN
XX 31-JUL-2003.
XX
XX 17-JAN-2003; 2003WO-US001981.
XX
XX 18-JAN-2002; 2002US-0350061P.
PR
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX

PI Huang F, Fairchild CR, Lee FY, Shaw P;
XX WPI; 2003-636735/60.
DR N-PSDB; ADD14760.
XX
XX New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.
XX
XX Claim 10; SEQ ID NO 346; 139pp; English.
XX
XX The present invention describes a predictor set comprising a plurality of
CC polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.
XX
SQ Sequence 101 AA;

Query Match 100.0%; Score 52; DB 7; Length 101;
Best Local Similarity 100.0%; Pred. No. 0.031;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACPLEKALDV 10
Db |||||
2 ACPLEKALDV 11

RESULT 11
ABM04788
ID ABM04788 standard; protein; 854 AA.
XX
AC ABM04788;
XX
DT 22-SEP-2003 (first entry)
XX
XX Human vacuolar assembly protein VSP41.
DE
XX spinal cord; neuropathic pain; central sensitisation pain; pain;
KW analgesic; gene therapy.
XX
XX Homo sapiens.
XX
XX EP1284298-A2.
PN
XX 19-FEB-2003.
XX
XX 26-JUL-2002; 2002EP-00255229.
XX
XX 27-JUL-2001; 2001GB-00018354.
PR
XX 07-FEB-2002; 2002GB-00002883.
PR

XX (WARN) WARNER LAMBERT CO.
 XX Brooksbank RA, Dixon AK, Lee K, Pinnock RD;
 XX WPI; 2003-543489/52.
 XX Use of an isolated gene sequence in the screening of compounds for
 PT diagnosing or treating pain.
 XX Claim 1; Page 45-48; 188pp; English.
 XX The invention relates to a novel isolated gene sequence that is
 CC downregulated in the spinal cord of a mammal in response to mechanically
 CC distinct first and second models of neuropathic or central sensitisation
 CC pain, useful in the screening of compounds for diagnosing or treating
 CC pain. A protein encoded by a gene of the invention has analgesic
 CC activity. A polynucleotide of the invention may have a use in gene
 CC therapy. The gene sequence is useful for preparing a composition for
 CC diagnosing or treating pain. The present sequence represents a protein
 CC encoded by a gene of the invention
 XX Sequence 854 AA;
 SQ

Query Match 82.7%; Score 43; DB 6; Length 854;
 Best Local Similarity 77.8%; Pred. No. 17;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
 DB 638 CPLEKALEI 646
 |||||:::
 638 CPLEKALEI 646

RESULT 12
 ABR62781
 ID ABR62781 standard; protein; 854 AA.
 XX AC ABR62781;
 XX 06-NOV-2003 (first entry)
 XX Human vascular protein sorting 41 (VPS41).
 XX VPS41; vascular protein sorting 41; human; apoptosis; cytostatic;
 KW gene therapy.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT Domain 571..712
 FT /note= "Clathrin domain"
 XX WO2003052070-A2.
 XX 26-JUN-2003.
 XX 12-DEC-2002; 2002WO-US039896.
 XX 13-DEC-2001; 2001US-0340309P.
 XX (EXEL-) EXELIXIS INC.
 XX Costa MA, Parry D, Chen C;
 XX WPI; 2003-533012/50.
 XX N-PSDB; ACF05818.
 XX Identifying a candidate cell death pathway modulating agent for
 PT diagnosing or treating cancer by detecting a test agent-biased activity
 PT of the assay system comprising a purified VPS41 polypeptide or nucleic
 XX acid.
 XX Claim 13; Page 42-45; 45pp; English.

XX The present sequence is the protein sequence of novel human vacuolar
 CC protein sorting 41 (VPS41). Genes have been discovered that increase
 CC germline apoptosis in *Caenorhabditis elegans*. The human homologue, VPS41,
 CC was subsequently identified. Expression analysis showed increased VPS41
 CC expression in 25% of head and neck cancers and 100% of skin cancers. The
 CC invention provides methods for using VPS41 polypeptide or nucleic acid to
 CC identify cell death pathway modulating agents. Such agents are candidate
 CC therapeutic agents for use in the treatment of disorders associated with
 CC defective or impaired cell death and/or VPS41 function. Preferred VPS41-
 CC modulating agents specifically bind to VPS41 polypeptides and restore
 CC cell death function. Others are nucleic acid modulators such as antisense
 CC oligonucleotides and RNAi, which repress VPS41 gene expression or product
 CC activity. A claimed method for diagnosing a disease utilizes a probe for
 CC VPS41 expression. The disease is cancer, especially head and neck cancer
 CC or skin cancer
 XX Sequence 854 AA;
 SQ

Query Match 82.7%; Score 43; DB 6; Length 854;
 Best Local Similarity 77.8%; Pred. No. 17;
 Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 CPLEKALDV 10
 DB 638 CPLEKALEI 646
 |||||:::
 638 CPLEKALEI 646

RESULT 13
 ABB59576
 ID ABB59576 standard; protein; 539 AA.
 XX AC ABB59576;
 XX 26-MAR-2002 (first entry)
 XX Drosophila melanogaster polypeptide SEQ ID NO 5520.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 XX WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US009231.
 XX 23-MAR-2000; 2000US-0191637P.
 XX 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 XX N-PSDB; ABL03679.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
 PT interactions.
 XX Disclosure; SEQ ID NO 5520; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 539 AA;

Query Match 75.0%; Score 39; DB 4; Length 539;
 Best Local Similarity 77.8%; Pred. No. 63;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 CPLEXALDV 10
 ||||: |||
 DB 288 CPLEELDV 296

RESULT 14
 AAE09797
 ID AAE09797 standard; protein; 751 AA.

XX AAE09797;

AC AAE09797;

XX 29-NOV-2001 (first entry)

DE Wheat oxidosqualene cyclase protein from clone wdk1c.pk010.o10.

XX Wheat; oxidosqualene cyclase; cytostatic; antilipemic; colon cancer;
 KW transgenic plant; insect tolerance; soy flour; bean; nut; tofu; natto;
 KW miso; tempeh.

XX Triticum aestivum.

XX WO200166773-A2.

PN 13-SEP-2001.

XX 09-MAR-2001; 2001WO-US007611.

XX 09-MAR-2000; 2000US-0188054P.

XX (DUPO) DU PONT DE NEMOURS & CO E I.

XX Bryan GT, McGonigle B, Maxwell CA, Potter SM, Hwang D;

XX WPI; 2001-565594/63.

DR N-PSDB; AAD16896.

XX New nucleic acid fragment encoding oxidosqualene cyclase, useful as DNA
 PT hybridization probe, polymerase chain reaction primer and for creating
 PT transgenic plants with cells having higher or lower of oxidosqualene
 PT cyclase.

XX Claim 12; Fig 1; 59pp; English.

XX The invention relates to oxidosqualene cyclase proteins and their
 CC corresponding cDNA molecules. Oxidosqualene cyclase genes are used as
 CC hybridization probes to screen libraries from any desired plant.
 CC Oxidosqualene cyclase DNA are used to create transgenic plants with cells
 CC having higher or lower level of oxidosqualene cyclase. These changes in
 CC the plant seed are useful to improve the seed nutritional value, and in
 CC the plant leaf may aid in insect tolerance. Oxidosqualene cyclase are
 CC produced in heterologous host cells and are used to prepare antibodies.
 CC The antibodies are useful for detecting oxidosqualene cyclase in situ in
 CC cells or in vitro in cell extracts. The invention is useful for altering
 CC level of expression of oxidosqualene cyclase (beta-amyrin synthase)
 CC protein in a plant cell. Soybean grain produced from a transgenic plant
 CC is useful for preparing food products such as soy protein product,
 CC soybean meal, soy flour, soy protein concentrate, soy milk, a dietary
 CC supplement, beans, nuts, tofu, natto, miso and tempeh. The formed soy
 CC protein product having increased saponin levels is useful for lowering
 CC cholesterol in a person suspected of having increased cholesterol, and
 CC for decreasing risk of colon cancer in a person suspected of having risk
 CC of cancer. The present sequence is wheat oxidosqualene cyclase protein

XX Sequence 751 AA;

Query Match 73.1%; Score 38; DB 4; Length 751;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 CPLEXALDV 10
 ||||: |||
 DB 363 CPINKALDM 371

RESULT 15

AAB72387

ID AAB72387 standard; protein; 101 AA.

XX AAB72387;

AC AAB72387;

XX 24-MAY-2001 (first entry)

DE Murine Mts 1 protein amino acid sequence.

XX Mts 1; Mts1-del75; Mts1-4S; S100A4; neuronal cell growth; cardiac arrest;
 KW neuronal degeneration; Parkinson's disease; Alzheimer's disease;
 KW Down's syndrome; stroke; sciatic crush; spinal cord injury; mouse.

XX Mus sp.

XX WO200118043-A2.

PN 15-MAR-2001.

XX 07-SEP-2000; 2000WO-US024495.

XX 10-SEP-1999; 99US-00393433.

XX (PROL-) PROLIFIA INC.

XX Bock E, Lukanidin EM, Berezin V;

XX WPI; 2001-235188/24.

XX New isolated functional derivatives of Mts protein for stimulating
 PT neuronal cell growth to treat neurological conditions e.g. Parkinson's
 PT disease, Alzheimer's disease, Down's syndrome and stroke.

XX Disclosure; Page; 60pp; English.

XX This invention relates to functional derivatives of an Mts protein,
 CC particularly Mts1-del75, and Mts1-4S. Mts 1 (also known as S100A4) is a
 CC member of the S100 gene family. Mts1 proteins and their functional
 CC derivatives are used to stimulate the growth of neuronal cells and in the
 CC treatment of neurological conditions including those characterised by
 CC neuronal degeneration, death or injury such as Parkinson's disease,
 CC Alzheimer's disease, Down's syndrome, stroke, cardiac arrest, sciatic
 CC crush, spinal cord injury, injury to sensory neurons or degenerative
 CC disease of the retina. The present sequence represents the murine Mts 1
 CC protein, from which the Mts proteins of the invention may be derived.
 CC NOTE: The present sequence is not shown in the specification but it has
 CC been derived from the sequence given in Genes Dev. 3, 1989 (referred to
 CC on page 10 of the specification)

XX Sequence 101 AA;

Query Match 69.2%; Score 36; DB 4; Length 101;
 Best Local Similarity 80.0%; Pred. No. 41;
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACPLEKALDV 10

DB 2 ARPLEALDV 11

Search completed: September 9, 2004, 10:53:03
 Job time : 24.0323 secs

us-10-067-618-4.rag

Thu Sep 9 12:03:21 2004

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:42:57 ; Search time 168.116 Seconds
(without alignments)
189.556 Million cell updates/sec

Title: US-10-067-618-2

Perfect score: 533

Sequence: 1 MACPLEKALDMVSTEHKYS.....IAMMCNEFFGPPDKQPRKK 101

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_xvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	ID	Description
1	508	95.3	101	6	Q9TV56
2	245	46.0	89	11	Q9R2B7
3	238	44.7	94	11	Q91V77
4	237	44.5	92	13	Q9PSF6
5	236	44.3	99	13	Q9AYJ2
6	234	43.9	94	11	Q9D3M4
7	232.5	43.6	55	6	Q9TSB1
8	230	43.2	92	11	Q925T3
9	201	37.7	100	13	Q7ZVA4
10	198	37.1	79	11	Q9JL08
11	185	34.7	101	13	Q93395
12	153	28.7	103	6	Q862H7
13	140.5	26.4	217	13	Q9YH57
14	134	25.1	77	6	Q862V4
15	131.5	24.7	124	11	Q9D708
16	130.5	24.5	213	4	Q9H4U1

17	130	24.4	80	11	Q91XG5
18	129.5	24.3	79	6	Q865V3
19	125	23.5	70	6	Q9TR16
20	121	22.7	148	11	Q8BLX1
21	121	22.7	2496	11	Q8VHD8
22	120.5	22.6	501	11	Q8CIU0
23	119.5	22.4	83	11	Q9QVR5
24	116.5	21.9	147	11	Q8C9X1
25	111.5	20.9	357	11	Q8BHC3
26	111.5	20.9	554	11	Q8K552
27	107	20.1	495	4	Q9UBG3
28	107	20.1	495	4	Q8N613
29	104.5	19.6	1218	4	Q05331
30	102.5	19.2	591	4	Q01720
31	102.5	19.2	687	4	Q9H4U2
32	95.5	17.9	225	5	Q86A72
33	94.5	17.7	73	6	Q28714
34	94.5	17.7	638	11	Q9D3P1
35	86.5	16.2	167	10	Q9AR93
36	83.5	15.7	136	5	Q86SC6
37	80.5	15.1	186	5	Q869M8
38	80	15.0	295	10	Q9MBG8
39	80	15.0	808	10	Q23052
40	79.5	14.9	349	3	Q8TFD0
41	79.5	14.9	349	3	Q8TGS1
42	79.5	14.9	349	3	Q8TFC4
43	78	14.6	525	5	Q86S94
44	77.5	14.5	2270	5	Q813E5
45	77	14.4	620	2	P77890

ALIGNMENTS

RESULT 1
Q9TV56 PRELIMINARY; PRT; 101 AA.
ID Q9TV56
AC Q9TV56;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Metastasin.
GN MTS1.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Madin-Darby; TISSUE=Kidney;
RA Miyamori H., Hasegawa K., Kim K., Sato H.;
RT "Expression of metastasin associated mts1 gene is co-induced with
RT membrane type-1 matrix metalloproteinase (MT1-MMP) during oncogenic
RT transformation and tubular formation of madin darby canine kidney
RT (MDCK) epithelial cells.";
RL Submitted (Aug-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AB031064; BAA83419.1; -.
DR HSSP; P30801; 1A03.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_s100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S100; 1.
DR ProDom; P0003407; CaBP_S100; 1.
DR ProDom; P0000012; EF-hand; 1.
DR PROSITE; PS000018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CaBP; 1.
SQ SEQUENCE 101 AA; 11847 MW; 5535387DB7577DF0 CRC64;
Query Match 95.3%; Score 508; DB 6; Length 101;
Best Local Similarity 95.0%; Pred. No. 6.9e-46;
Matches 96; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Q91XG5 mus musculus
Q865V3 equs cabal
Q9TR16 bos taurus
Q8BLX1 mus musculus
Q8VHD8 mus musculus
Q8CIU0 rattus norv
Q9QVR5 mus musculus
Q8C9X1 mus musculus
Q8BHC3 mus musculus
Q8K552 mus musculus
Q8UBG3 homo sapien
Q8N613 homo sapien
Q05331 homo sapien
Q01720 homo sapien
Q9H4U2 homo sapien
Q86A72 dictyosteli
Q28714 oryctolagus
Q9D3P1 mus musculus
Q9AR93 medicago sa
Q86SC6 ciona intes
Q869M8 dictyosteli
Q9MBG8 arabidopsis
Q23052 arabidopsis
Q8TFD0 saccharomyc
Q8TGS1 saccharomyc
Q8TFC4 saccharomyc
Q86S94 babesia rod
Q813E5 plasmodium
P77890 legionella

```

QY 1 MACPLEKALDVMVSTFKYSGKGDGKFLNKLSELKELLTRELPSFLGKRTDEAAFOKLMS 60
DB 1 MTEPLEKALDVMVSTFKYSGKGDGKFLNKLSELKELLTRELPSFLGKRTDEAAFOKLMS 60

QY 61 NLDSNRDNEVDFOEYCVFLSCIAMCNEFFEGPDKOPRKK 101
DB 61 NLDSNRDNEVDFOEYCVFLSCIAMCNEFFEGPDKOPRKK 101

RESULT 2
Q9R2B7 PRELIMINARY; PRT; 89 AA.
AC Q9R2B7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Calcyclin (Calcium binding protein).
GN CACY OR S100A6.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RA Konrad L., Gabius H.J., Ammuller G.;
RT "Sequence and expression study of calcyclin in the rat testis.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar;
RA Ito M., Kizawa K.;
RT "Expression of S100 genes in hair follicle epithelium.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
-1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
CC EMBL; AJ132717; CAB42002.1; -
DR EMBL; AF140232; AAK28306.1; -
DR HSSP; P30801; 1A03.
DR GO; GO:0005635; C:nuclear membrane; ISS.
DR GO; GO:0005509; F:calcium ion binding; ISS.
DR GO; GO:0005515; F:protein binding; ISS.
DR GO; GO:0007409; P:axonogenesis; ISS.
DR GO; GO:0048146; P:positive regulation of fibroblast prolifera. . .; ISS.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProSite; PS00018; EF_HAND; 1.
DR ProSite; PS00303; S100_CABP; 1.
KW Cyclin.
SQ SEQUENCE 89 AA; 10035 MW; 2AA1A4163D57DC87 CRC64;

Query Match 46.0%; Score 245; DB 11; Length 89;
Best Local Similarity 52.3%; Pred. No. 2.9e-18;
Matches 46; Conservative 19; Mismatches 21; Indels 2; Gaps 1;

QY 1 MACPLEKALDVMVSTFKYSGKGDGKFLNKLSELKELLTRELPSFLGKRTDEAAFOKLMS 60
DB 1 MACPLDQALGLLVAFHKYSGKGDGKHTLSKELKELIQKELT--IGAKLQDAETARLMD 58

QY 61 NLDSNRDNEVDFOEYCVFLSCIAMCNE 88
DB 59 DLDNRKQDVNFQYVAFGLALALYNE 86

RESULT 3
Q91V77 PRELIMINARY; PRT; 94 AA.
AC Q91V77;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

```

```

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 11 days embryo cDNA, RIKEN full-length enriched library,
DE clone:2700088D09, full insert sequence (S100 calcium binding protein
DE A1) (Adult male kidney cDNA, RIKEN full-length enriched library,
DE clone:0610031F03, full insert sequence).
GN S100A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA Arakawa T., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Inotani K., Ishii Y., Itoh M., Izawa M., Kato H., Kawai J., Kojima Y.,
RA Konno H., Kouda M., Koya S., Kurihara C., Mateyama T., Miyazaki A.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T.,
RA Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D.,
RA Shibata K., Shibata Y., Shinagawa A., Shiraki T., Sogabe Y.,
RA Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y.,
RA Toya T., Yamamuro T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA MEDLINE=21085660; PubMed=11217851;
RX RIKEN FANTOM Consortium.;
RA "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA MEDLINE=99279253; PubMed=10349636;
RX Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
[4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA MEDLINE=20499374; PubMed=11042159;
RX Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
[5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney;
RA MEDLINE=20530913; PubMed=11076861;
RX Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
[6]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
[7]
RP SEQUENCE FROM N.A.
RC STRAIN=129/Sv;
RA Du X.-J., Cole T.J., Tennis N., Gao X.-M., Kontgen F., Kemp B.E.,
RA Heierhorst J.;
RT "Impaired cardiac contractility without cardiomyopathy in S100A1-
RT deficient mice.";

```

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.

DR EMBL; AK012578; BAB28330.1; -.
DR EMBL; BC005590; AAO05590.1; -.
DR EMBL; AF368423; AAL14436.1; -.
DR EMBL; AK002721; BAB22308.1; -.
DR MGD; MGI:1338917; S100a1.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR Pfam; PF00036; EF-hand.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF-HAND; 1.
DR PROSITE; PS00303; S100 CABP; 1.
DR PROSITE; PS00303; S100 CABP; 1.
SQ SEQUENCE 94 AA; 10505 MW; 7C8E817BBF36ED15 CRC64;

Query Match 44.7%; Score 238; DB 11; Length 94;
Best Local Similarity 48.9%; Pred. No. 1.7e-17;
Matches 44; Conservative 16; Mismatches 30; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMVSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60

Db 1 MGSELESAMETLNVFHAHSGEGDKYKLSKKELKDLQLTGLSGFLDVQKDAVDKVMK 60

QY 61 NLDSNRDNEVDFOEYCVFLSCIAMMCNEFF 90

Db 61 ELDENGGEYDFKEYVYVLAALTVACNFF 90

RESULT 4

Q9PSF6 ID Q9PSF6 PRELIMINARY; PRT; 92 AA.

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Ictaluridin.

OS Ictalurus punctatus (Channel catfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Siluriformes;

OC Ictaluridae; Ictalurus.

OX NCBI_TaxID=7998;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94376615; PubMed=8090068;

RA Bettini E., Porta A.R., Dahmen N., Wang H., Margolis F.I.;

RT "Expressed sequence tags (EST) identify genes preferentially expressed

in catfish chemosensory tissues";

RL Brain Res. Mol. Brain Res. 23:285-291(1994).

CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.

DR HSP; P30801; 1A03.

DR GO; GO:0005509; F:calcium ion binding; IEA.

DR InterPro; IPR001751; CaBP_S100.

DR InterPro; IPR002048; EF-hand.

DR Pfam; PF00036; EF-hand; 1.

DR Pfam; PF01023; S_100; 1.

DR ProDom; PD003407; CaBP_S100; 1.

DR ProDom; PD000012; EF-hand; 1.

DR PROSITE; PS00018; EF-HAND; 1.

DR PROSITE; PS00303; S100 CABP; 1.

SQ SEQUENCE 92 AA; 10022 MW; E04875D0C9921C50 CRC64;

Query Match 44.5%; Score 237; DB 13; Length 92;
Best Local Similarity 50.0%; Pred. No. 2.1e-17;
Matches 42; Conservative 17; Mismatches 25; Indels 0; Gaps 0;

QY 5 LEKALDVMVSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 64

Db 4 LQGMALLSTFKYSGEGDKTLTKGELKDLTLTKELGAFGNCSDQATLDFKDLDT 63

QY 65 NRDNEDVDFOEYCVFLSCIAMMCNE 88

Db 64 NADGVDFQFYATWVACTTMLCNK 87

RESULT 5

Q8AYJ2 ID Q8AYJ2 PRELIMINARY; PRT; 99 AA.

AC Q8AYJ2; 2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

OS S-100 calcium-binding protein A1.

OS Squalus acanthias (Spiny dogfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;

OC Elasmobranchii; Squala; Squalidae; Squalus.

OX NCBI_TaxID=7797;

RN [1]

RP SEQUENCE FROM N.A.

RA Wang C., Callard G.V.;

RT "Molecular cloning and stage dependence of an S-10 cDNA from the shark

testis";

RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF421551; AAN63527.1; -.

GO; GO:0005509; F:calcium ion binding; IEA.

DR InterPro; IPR001751; CaBP_S100.

DR InterPro; IPR002048; EF-hand.

DR Pfam; PF00036; EF-hand; 1.

DR Pfam; PF01023; S_100; 1.

DR ProDom; PD003407; CaBP_S100; 1.

DR ProDom; PD000012; EF-hand; 1.

DR SMART; SM00054; EPH; 1.

DR PROSITE; PS00018; EF-HAND; 1.

DR PROSITE; PS00303; S100 CABP; 1.

SQ SEQUENCE 99 AA; 11050 MW; BA62D8D190A4A3693 CRC64;

Query Match 44.3%; Score 236; DB 13; Length 99;
Best Local Similarity 48.8%; Pred. No. 2.9e-17;
Matches 42; Conservative 20; Mismatches 24; Indels 0; Gaps 0;

QY 5 LEKALDVMVSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 64

Db 4 LESAVAGIIVFPRKYSGEGDKYSLNNEVDLLKAEPLNFKLSQDKAAVDKIMKDLDR 63

QY 65 NRDNEDVDFOEYCVFLSCIAMMCNEFF 90

Db 64 NKDGELDFQEFVVLIAALAAACNDF 89

RESULT 6

Q9D3M4 ID Q9D3M4 PRELIMINARY; PRT; 94 AA.

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE S100 calcium binding protein A1.

GN S100A1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=C57BL/6J; TISSUE=Head;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

Arakawa T., Harai A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamada I.,

Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,

Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

```
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AK017279; BAB30670.1; -.
DR HSSP; P04631; 1B4C.
DR MGD; MGI:1338917; Sl00a1.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100 CABP; 1.
SQ SEQUENCE 94 AA; 10487 MW; 7C9E817D135C2D15 CRC64;

Query Match 43.9%; Score 234; DB 11; Length 94;
Best Local Similarity 47.8%; Pred. No. 4.5e-17;
Matches 43; Conservative 17; Mismatches 30; Indels 0; Gaps 0;

QY 1 MACPLEKALDVWVSTFHKYSKGGKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
DB 1 MGSELSAMELIINVFFHSHSGKGGYKLSKELKDLLQTELSGLDVGKDADVDKVIK 60

QY 61 NLDSNRDNEVDFOBYCVFLSCIAMCNFF 90
DB 61 ELDENGDGEVDKEYVVLVAALTACNFF 90

RESULT 7
ID Q9TSB1 PRELIMINARY; PRT; 55 AA.
AC Q9TSB1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Calvasculin (Fragments).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RX MEDLINE=92117670; PubMed=1731618;
RA Watanabe Y., Kobayashi R., Ishikawa T., Hidaka H.;
RT "Isolation and characterization of a calcium-binding protein derived
RT from mRNA termed pKa, pEL-98, 18A2, or 42A by the newly synthesized
RT vasorelaxant W-66 affinity chromatography.";
RL Arch. Biochem. Biophys. 292:563-569(1992).
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR PIR; S20330; S20330.
DR HSSP; P30801; 1A03.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR ProDom; PD003407; CaBP_S100; 1.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100 CABP; 1.
FT NON_TER 1
FT NON_CONS 11
FT NON_TER 55

SQ SEQUENCE 55 AA; 6452 MW; 878498A4BAA8BD86 CRC64;

Query Match 43.6%; Score 232.5; DB 6; Length 55;
Best Local Similarity 69.4%; Pred. No. 3.5e-17;
Matches 50; Conservative 1; Mismatches 4; Indels 17; Gaps 1;

QY 8 ALDVWVSTFHKYSKGGKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMSNLDSNRD 67
DB 1 ALDVWVSTFH-----KELLTRELPSFLGKXTDETAFQKLMSNLDSYKND 43

QY 68 NEVDFOEYCVFL 79
DB 44 NEVDFOEYKXVFL 55

RESULT 8
ID Q925T3 PRELIMINARY; PRT; 92 AA.
AC Q925T3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S100B.
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RA Kuge O., Yamakawa Y., Nishijima M.;
RT "Enhancement of transport-dependent decarboxylation of
RT phosphatidylserine by S100B protein in permeabilized Chinese hamster
RT ovary cells.";
RL J. Biol. Chem. 276:10000-10004(2001).
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AB056121; BAB3945.1; -.
DR GO; GO:0005737; C:cytoplasm; ISS.
DR GO; GO:0005576; C:extracellular; ISS.
DR GO; GO:0005509; F:calcium ion binding; ISS.
DR GO; GO:0012110; F:kinase inhibitor activity; ISS.
DR GO; GO:0042803; F:protein homodimerization activity; ISS.
DR GO; GO:0048155; F:S100 alpha binding; ISS.
DR GO; GO:0048154; F:S100 beta binding; ISS.
DR GO; GO:0048156; F:tau protein binding; ISS.
DR GO; GO:0008270; F:zinc ion binding; ISS.
DR GO; GO:0008143; P:astrocyte activation; ISS.
DR GO; GO:0007409; P:axonogenesis; ISS.
DR GO; GO:0006874; P:calcium ion homeostasis; ISS.
DR GO; GO:0006112; P:energy reserve metabolism; ISS.
DR GO; GO:0048151; P:hyperphosphorylation; ISS.
DR GO; GO:0006917; P:induction of apoptosis; ISS.
DR GO; GO:0007611; P:learning and/or memory; ISS.
DR GO; GO:0045917; P:positive regulation of complement activation; ISS.
DR GO; GO:0042035; P:regulation of cytokine biosynthesis; ISS.
DR GO; GO:0048169; P:regulation of long-term neuronal synaptic p. .; ISS.
DR GO; GO:0006417; P:regulation of protein biosynthesis; ISS.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100 CABP; 1.
SQ SEQUENCE 92 AA; 10749 MW; AF50107EC2BEDF6B CRC64;

Query Match 43.2%; Score 230; DB 11; Length 92;
Best Local Similarity 49.4%; Pred. No. 1.2e-16;
Matches 43; Conservative 15; Mismatches 29; Indels 0; Gaps 0;

QY 5 LEKALDVWVSTFHKYSKGGKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMSNLDS 64
DB 4 LEKAWALIDIFHQYSGREGDKHLKKSELKLNELSHFLFEIKFQEVVDKVMETLDE 63
```

QY 65 NRNEVDFOEYCVFVLSCIAMMNEFFE 91
Db 64 DGGECDFQEFMAFVSMVTACHEFFE 90

RESULT 9
Q7ZVA4 PRELIMINARY; PRT; 100 AA.
AC Q7ZVA4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Body;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC045941; AAH45941.1; -.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR SMART; SM00054; EFh_1-
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Hypothetical protein.
SQ SEQUENCE 100 AA; 11157 MW; DAB81814E54CB8C8 CRC64;

Query Match 37.7%; Score 201; DB 13; Length 100;
Best Local Similarity 38.5%; Pred. No. 1.4e-13;
Matches 35; Conservative 27; Mismatches 29; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTSTFKYSGKGDGKFLNKSSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db 1 MPSDLERAMETLTVFHYRSGAGNSLTSLRELKQLMEKELASFLKSQKDPRAVDXIMK 60

QY 61 NLDNRDNEVDFOEYCVFVLSCIAMMNEFFE 91
Db 61 DLDANGDGEVNFEEVSLVVGSLTACRQLYQ 91

RESULT 10
Q9JL08 PRELIMINARY; PRT; 79 AA.
AC Q9JL08;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S100 calcium binding protein A1 (Fragment).
GN S100A1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Kiewitz R., Lyons G.E., Schafer B.W., Heizmann C.W.;
RT "Transcriptional regulation of S100a1 and expression during mouse heart development";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AF218353; AAF32320.1; -.
DR HSSP; P04631; 1B4C.
DR MGD; MGI:1338917; S100a1.

DR GO: GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
FT NON_TER 79
SQ SEQUENCE 79 AA; 8863 MW; F94EDA3A798815D5 CRC64;

Query Match 37.1%; Score 198; DB 11; Length 79;
Best Local Similarity 48.1%; Pred. No. 2.3e-13;
Matches 38; Conservative 14; Mismatches 27; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTSTFKYSGKGDGKFLNKSSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db 1 MGSELESAMETLNVFHAHSAQEGDKYKLSKELKDLQLQTELSGFLDQKDAVAVKMK 60

QY 61 NLDNRDNEVDFOEYCVFVLS 79
Db 61 ELDENGSDGEVDFKEYVVLV 79

RESULT 11
O93395 PRELIMINARY; PRT; 101 AA.
AC O93395;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE S100-like calcium binding protein.
GN S100.
OS Salvelinus fontinalis (Brook trout) (Brook char).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Salvelinus.
OX NCBI_TaxID=8038;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovary; ovary;
RX MEDLINE=20534789; PubMed=11080585;
RA Robe J., Gostz F.W.;
RT "A S100 homologue mRNA isolated by differential display PCR is down-regulated in the brook trout (Salvelinus fontinalis) post-ovulatory ovary.";
RL Gene 257:187-194(2000).
CC -1- SIMILARITY: BELONGS TO THE S-100 FAMILY.
DR EMBL; AF077613; AAC28367.1; -.
DR HSSP; P02633; 4ICB.
DR GO: GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
SQ SEQUENCE 101 AA; 11285 MW; BF45582FF9279D0A CRC64;

Query Match 34.7%; Score 185; DB 13; Length 101;
Best Local Similarity 35.6%; Pred. No. 7e-12;
Matches 36; Conservative 24; Mismatches 41; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTSTFKYSGKGDGKFLNKSSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db 1 MPSQLESSMESLTVFHYRVDKGDGDCNTLSKKELMQTELASFLKSQKDPAAIDTIMK 60

QY 61 NLDNRDNEVDFOEYCVFVLSCIAMMNEFFGPFDPKPRKK 101
Db 61 DLDQNGDGKVSFEFVSLVVGSLVGLSIAEQIYQLHTKKVAKK 101

```
RESULT 12
Q862H7
ID Q862H7 PRELIMINARY; PRT; 103 AA.
AC Q862H7;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to S100 calcium-binding protein A11 (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22544902; PubMed=12658628;
RA Ishiwa H., Katsuma S., Kizaki K., Patel O.V., Nakano H.,
RA Takahashi T., Imai K., Hirasawa A., Shiojima S., Ikawa H., Suzuki Y.,
RA Tsujimoto G., Izaik Y., Todoroki J., Hashizume K.;
RT "Characterization of gene expression profiles in early bovine
pregnancy using a custom cDNA microarray.";
RL Mol. Reprod. Dev. 65:9-18(2003).
DR EMBL; AB099012; BAC56502.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efhand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR SMART; SM00054; EFh; 1.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
FT NON_TER 1
FT NON_TER 103
SQ SEQUENCE 103 AA; 11382 MW; 1C8DFE7716384281 CRC64;

Query Match 28.7%; Score 153; DB 6; Length 103;
Best Local Similarity 30.6%; Pred. No. 1.7e-08;
Matches 26; Conservative 25; Mismatches 34; Indels 0; Gaps 0;

QY 6 EKALDVMTSTHKYSGKDGKFKLNKSELKELLTRELPSFLGKRTDEAFAFKLMNLSN 65
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 11 ERCIESLAVFQKHAGRDGNNKLSKABFLFMNTELGFTKNQKDPGLDRMMKKLDIN 70
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 66 RDNEVDFOEYCVFLSCIAMMCNERF 90
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 71 SDGQLDFQEFNLIGGLAIACHESF 95
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 13
Q9YH57
ID Q9YH57 PRELIMINARY; PRT; 217 AA.
AC Q9YH57;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE P2601f.
GN P2601f.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory organ;
RX MEDLINE=99008915; PubMed=9791000;
RA Miwa N., Kobayashi M., Takamatsu K., Kawamura S.;
RT "Purification and molecular cloning of a novel calcium-binding
protein, p2601f, in the frog olfactory epithelium.";
RL Biochem. Biophys. Res. Commun. 251:860-867(1998).
CC -!- SIMILARITY: BELONGS TO THE S-100 FAMILY.

DR EMBL; D50333; BAA34388.1; -.
DR PIR; JEO330; JEO330.
DR HSP; P31950; 1QLS.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF01023; S_100; 2.
DR ProDom; PD003407; CaBP_S100; 2.
SQ SEQUENCE 217 AA; 24494 MW; 0403689A9543810B CRC64;

Query Match 26.4%; Score 140.5; DB 13; Length 217;
Best Local Similarity 28.6%; Pred. No. 8.1e-07;
Matches 28; Conservative 28; Mismatches 37; Indels 5; Gaps 2;

QY 5 LEKALDVMTSTHKYSGKDGKFKLNKSELKELLTRELPSFLGKRTD-EAFAFKLMN-- 61
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 12 MERSNEKIITVFQRYAGKEGNTTSNNFKEFQPMNAELGSGFTKNQKDPAILRKIMKSV 71
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 --LOSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQ 97
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 72 GGVGDKQDGLDFQEFNLIGGMVAVACHEALMKCPPTQ 109
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 14
Q862V4
ID Q862V4 PRELIMINARY; PRT; 77 AA.
AC Q862V4;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Similar to calpactin I light chain (Fragment).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22544902; PubMed=12658628;
RA Ishiwa H., Katsuma S., Kizaki K., Patel O.V., Nakano H.,
RA Takahashi T., Imai K., Hirasawa A., Shiojima S., Ikawa H., Suzuki Y.,
RA Tsujimoto G., Izaik Y., Todoroki J., Hashizume K.;
RT "Characterization of gene expression profiles in early bovine
pregnancy using a custom cDNA microarray.";
RL Mol. Reprod. Dev. 65:9-18(2003).
DR EMBL; AB098817; BAC56353.1; -.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR ProDom; PD003407; CaBP_S100; 1.
DR PROSITE; PS00303; S100_CABP; 1.
FT NON_TER 1
FT NON_TER 77
SQ SEQUENCE 77 AA; 8788 MW; D057B286A0132750 CRC64;

Query Match 25.1%; Score 134; DB 6; Length 77;
Best Local Similarity 37.3%; Pred. No. 1.2e-06;
Matches 25; Conservative 15; Mismatches 27; Indels 0; Gaps 0;

QY 24 GDKFKLNKSELKELLTRELPSFLGKRTDEAFAFKLMNLSNDRNEVDFOEYCVFLSCIA 83
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 GDKGYLTKEDLRLVLMKEFFGLFNQKDPGLAVDKIMKLDQCRDGKVGFSFSLIAGLT 60
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 84 MMCNEFF 90
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 IACNDYF 67
| : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
Q9D708
ID Q9D708 PRELIMINARY; PRT; 124 AA.
AC Q9D708;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
```


Search completed: September 9, 2004, 10:58:13
Job time : 170.116 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:32:57 ; Search time 28.671 Seconds
(without alignments)
183.429 Million cell updates/sec

Title: US-10-067-618-2

Perfect score: 533

Sequence: 1 MACPLEKALDVMVSTFKYS.....IAMNCNEFFGFPDKQPRKK 101

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	533	100.0	101	1 S104_HUMAN	P26447 homo sapien
2	505	94.7	100	1 S104_BOVIN	P35466 bos taurus
3	496	93.1	101	1 S104_MOUSE	P07091 mus musculus
4	488	91.6	101	1 S104_RAT	P05942 rattus norv
5	336	63.0	97	1 S102_BOVIN	P10462 bos taurus
6	323	60.6	97	1 S102_HUMAN	P29034 homo sapien
7	250	46.9	90	1 S106_RAT	P05964 rattus norv
8	248	46.5	90	1 S106_HUMAN	P06703 homo sapien
9	246	46.2	89	1 S106_MOUSE	P14069 mus musculus
10	244	45.8	90	1 S106_RABIT	P30801 oryctolagus
11	242	45.4	93	1 S10A_BOVIN	P02639 bos taurus
12	241	45.2	92	1 S101 ICTPU	Q91061 ictalurus p
13	241	45.2	93	1 S10A_HUMAN	P23297 homo sapien
14	239	44.8	92	1 S106_CHICK	Q98953 gallus gall
15	239	44.8	93	1 S105_MOUSE	O88945 mus musculus
16	239	44.8	93	1 S10A_RAT	P35467 rattus norv
17	233.5	43.8	92	1 S105_HUMAN	P33763 homo sapien
18	232	43.5	91	1 S10B_BOVIN	P02638 bos taurus
19	231	43.3	91	1 S10B_RAT	P04631 rattus norv
20	231	43.3	93	1 S10A_MOUSE	P56565 mus musculus
21	229	43.0	91	1 S10B_HUMAN	P04271 homo sapien
22	229	43.0	92	1 S106_HORSE	O77691 equus cabal
23	229	43.0	101	1 S103_HUMAN	P33764 homo sapien
24	228	42.8	91	1 S10B_MOUSE	P50114 mus musculus
25	223	41.8	101	1 S103_MOUSE	Q8WKG8 homo sapien
26	222	41.7	98	1 S102_HUMAN	P27003 gallus gall
27	200	37.5	95	1 S10P_HUMAN	P25915 homo sapien
28	172.5	32.4	96	1 S101_CHICK	P24479 gallus gall
29	170	31.9	101	1 S111_CHICK	P08207 mus musculus
30	169.5	31.8	96	1 S111_MOUSE	P04163 sus scrofa
31	168.5	31.6	95	1 S110_PIG	P08206 homo sapien
32	168.5	31.6	96	1 S110_HUMAN	P50116 rattus norv
33	165.5	31.1	112	1 S109_RAT	

34	163.5	30.7	94	1 S110_RAT	P05943 rattus norv
35	160	30.0	98	1 S111_MOUSE	P50543 mus musculus
36	160	30.0	119	1 M126_CHICK	P28318 gallus gall
37	159.5	29.9	96	1 S110_XENLA	P27004 xenopus lae
38	159.5	29.9	105	1 S111_HUMAN	P31949 homo sapien
39	158	29.6	102	1 S111_RABIT	P24480 oryctolagus
40	157.5	29.5	91	1 S112_HUMAN	P80511 homo sapien
41	156.5	29.4	112	1 S109_MOUSE	P31725 mus musculus
42	156	29.3	91	1 S112_PIG	P80310 sus scrofa
43	153	28.7	91	1 S112_BOVIN	P79105 bos taurus
44	150.5	28.2	114	1 S109_HUMAN	P06702 homo sapien
45	145.5	27.3	98	1 S113_MOUSE	P97352 mus musculus

ALIGNMENTS

RESULT 1
S104_HUMAN
ID S104_HUMAN STANDARD; PRT; 101 AA.
AC P26447;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Placental calcium-binding protein (Calvasculin) (S100 calcium-binding protein A4) (MTS1 protein).
DE S100A4 OR CAPL OR MTS1.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=93041710; PubMed=1384693;
RA Engelkamp D., Schaefer B.W., Erne P., Heizmann C.W.;
RT "S100 alpha, CAPL, and CACY: molecular cloning and expression analysis of three calcium-binding proteins from human heart.";
RT Biochemistry 31:10258-10264(1992).
RL [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=93342029; PubMed=8341667;
RX Engelkamp D., Schaefer B., Mattei M.-G., Erne P., Heizmann C.W.;
RT "Six S100 genes are clustered on human chromosome 1q21: identification of two genes coding for the two previously unreported calcium-binding proteins S100D and S100E.";
RT Proc. Natl. Acad. Sci. U.S.A. 90:6547-6551(1993).
RL [3]
RN SEQUENCE FROM N.A.
RP TISSUE=Spleen;
RX MEDLINE=93028421; PubMed=1329089;
RA Tsuchinsky E.M., Ford H.L., Kramerov D., Reshetnyak E., Grigorian M., Zain S., Lukanidin E.;
RT "Transcriptional analysis of the mts1 gene with specific reference to 5' flanking sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 89:9146-9150(1992).
RL [4]
RN SEQUENCE FROM N.A.
RP TISSUE=Cervix, and Prostate;
RX MEDLINE=22388257; PubMed=12477932;
RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko M., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton L., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loguillano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Heiton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

```

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M80563; AAA51920.1; -.
CC DR EMBL; Z18950; CAA79474.1; -.
CC DR EMBL; Z33457; CAA83880.1; -.
CC DR EMBL; BC000838; AAH00838.1; -.
CC DR EMBL; BC016300; AAH16300.1; -.
CC DR PIR; A48219; A48219
CC DR PDB; 1M31; 30-OCT-02.
CC DR SWISS-2DPAGE; P26447; HUMAN.
CC DR Genew; HGNC:10494; S100A4.
CC DR MIM; 114210; -.
CC DR GO; GO:0005509; F:calcium ion binding; TAS.
CC DR InterPro; IPR001751; CaBP_S100.
CC DR InterPro; IPR002048; EF-hand.
CC DR Pfam; PF00036; ehand; 1.
CC DR Pfam; PF01023; S100; 1.
CC DR ProDom; PD003407; CaBP_S100; 1.
CC DR ProDom; PD000012; EF-hand; 1.
CC DR PROSITE; PS00018; EF_HAND; 1.
CC DR PROSITE; PS00303; S100_CABP; 1.
CC DR Calcium-binding; 3D-structure.
CC FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
CC FT CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
CC FT SEQUENCE 101 AA; 11728 MW; 286D2B7B07EDB562 CRC64;
SQ
Query Match 100.0%; Score 533; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.6e-43;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MACPLEKALDVNVSTPHKYSKGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
DB 1 MACPLEKALDVNVSTPHKYSKGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
QY 61 NLDSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101
DB 61 NLDSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101
RESULT 2
S104_BOVIN STANDARD; PRT; 100 AA.
AC P35466;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Placental calcium-binding protein homolog.
GN S100A4 OR CAPL.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE.
RC TISSUE=Retina;

```

```

RX MEDLINE=94164991; PubMed=8119967;
RA Polans A.S., Palczewski K., Asson-Batres M.A., Ohguro H., Witowska D.,
RA Haley T.L., Baizer L., Crabb J.W.;
RT "Purification and primary structure of Capl, an S-100-related
RT calcium-binding protein isolated from bovine retina."
RL J. Biol. Chem. 269:6233-6240 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98430466; PubMed=9759666;
RA Duarte W.R., Kasugai S., Iimura T., Oida S., Takenaga K., Ohya K.,
RA Ishikawa I.;
RT "cDNA cloning of S100 calcium-binding proteins from bovine
RT periodontal ligament and their expression in oral tissues."
RL J. Dent. Res. 77:1694-1699 (1998).
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D89056; BAA13754.1; -.
CC DR PIR; A53217; A53217.
CC DR HSSP; P30801; 1A03.
CC DR InterPro; IPR001751; CaBP_S100.
CC DR InterPro; IPR002048; EF-hand.
CC DR Pfam; PF00036; ehand; 1.
CC DR Pfam; PF01023; S100; 1.
CC DR ProDom; PD003407; CaBP_S100; 1.
CC DR ProDom; PD000012; EF-hand; 1.
CC DR PROSITE; PS00018; EF_HAND; 1.
CC DR PROSITE; PS00303; S100_CABP; 1.
CC DR Calcium-binding; Acetylation.
CC KW INIT MET 0
CC FT MOD_RES 1 1 ACETYLATION.
CC FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
CC FT CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
CC FT SEQUENCE 100 AA; 11675 MW; DFFCAA7561DSEFA4 CRC64;
SQ
Query Match 94.7%; Score 505; DB 1; Length 100;
Best Local Similarity 96.0%; Pred. No. 6.9e-41;
Matches 96; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 2 ACPLKALDVNVSTPHKYSKGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMSN 61
DB 1 AYPLEKALDVNVSTPHKYSKGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMSN 60
QY 62 LQSNRDNEDVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101
DB 61 LQCNKDNEDVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 100
RESULT 3
S104_MOUSE STANDARD; PRT; 101 AA.
AC P07091; P20066;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Placental calcium-binding protein (18A2) (PBL98) (MTS1 protein)
DE (Metastatic cell protein).
GN S100A4 OR CAPL OR MTS1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/C;

```

RX MEDLINE=87316927; PubMed=3628004;
RA Jackson-Grusby L.L., Swiergiel J., Linzer D.I.H.;
RT "A growth-related mRNA in cultured mouse cells encodes a placental
RL calcium binding protein.";
RL Nucleic Acids Res. 15:6677-6690(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BAUB/C;
RX MEDLINE=88198109; PubMed=3162911;
RA Goto K., Endo H., Fujiyoshi T.;
RT "Cloning of the sequences expressed abundantly in established cell
RL lines: identification of a cDNA clone highly homologous to S-100, a
RL calcium binding protein.";
RN J. Biochem. 103:48-53(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=89378739; PubMed=2550322;
RA Ebralidze A., Tulchinsky E., Grigorian M., Afanasyeva A., Senin V.,
RA Revazova E., Lukanidin E.;
RT "Isolation and characterization of a gene specifically expressed in
RT different metastatic cells and whose deduced gene product has a high
RL degree of homology to a Ca2+-binding protein family.";
RN Genes Dev. 3:1086-1093(1989).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=90236313; PubMed=2332170;
RA Tulchinsky E.M., Grigorian M.S., Ebralidze A.K., Milshina N.I.,
RA Lukanidin E.M.;
RT "Structure of gene mts1, transcribed in metastatic mouse tumor
RL cells.";
RN Gene 87:219-223(1990).
RN [5]
RP SEQUENCE OF 1-54 FROM N.A.
RX MEDLINE=93141279; PubMed=8423998;
RA Tulchinsky E., Kramerov D., Ford H.L., Reshetnyak E., Lukanidin E.,
RA Zain S.;
RT "Characterization of a positive regulatory element in the mts1 gene.";
RN Oncogene 8:79-86(1993).
CC -1- TISSUE SPECIFICITY: Specifically expressed in different metastatic
CC cells.
CC -1- INDUCTION: The mRNA coding for this protein increases in
CC abundance after serum stimulation of quiescent mouse fibroblasts.
CC -1- SIMILARITY: Belongs to the S-100 family.
CC -1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X05835; CAA29282.1; -;
CC EMBL; D00208; BAA00148.1; -;
CC EMBL; M36578; AAA39749.1; -;
CC EMBL; M36579; AAA39750.1; -;
CC EMBL; X16190; CAA34316.1; -;
CC EMBL; X16094; CAA34224.1; -;
CC PIR; S06207; S06207.
CC HSP; P30801; IAO3.
CC MGD; MGI:1330282; S100a4.
CC InterPro; IPR001751; CaBP_S100.
CC InterPro; IPR002048; EF-hand.
CC Pfam; PF00036; efhand; 1.
CC ProDom; PD003407; CaBP_S100; 1.
CC PROSITE; PS00018; EF_HAND; 1.
CC PROSITE; PS00303; S100_CaBP; 1.
CC KX Calcium-binding; Placenta.
CC CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
CC CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
CC FT

FT CONFLICT 47 47 G -> GVSGSXFNGQ (IN REF. 5).
SQ SEQUENCE 101 AA; 11721 MW; 2302254B67A4C873 CRC64;
Query Match 93.1%; Score 496; DB 1; Length 101;
Best Local Similarity 93.1%; Pred. No. 4.9e-40;
Matches 94; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY 1 MACPLEKALDVNVSTFHKYSGKEGDKFLKNSKELLTRELPSFLGKRTDEAAAFQKWS 60
DB 1 MARPLEEALDIVSTFHKYSGKEGDKFLKNTKELLTRELPSFLGKRTDEAAAFQKWS 60
QY 61 NLDNRDNEVDFOEYCVFLSCIAMMNCNEFFEGFPDQPRKK 101
DB 61 NLDNRDNEVDFOEYCVFLSCIAMMNCNEFFEGFPDQPRKK 101
RESULT 4
ID S104 RAT STANDARD; PRT; 101 AA.
AC P05942;
DT 01-NOV-1988 (Rel. 09, Created)
DT 11-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Placental calcium-binding protein (Nerve growth factor induced protein
DE 42a) (P9K).
GN S100A4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88125019; PubMed=3422491;
RA Masiakowski P., Shooter E.M.;
RT "Nerve growth factor induces the genes for two proteins related to a
RL family of calcium-binding proteins in PC12 cells.";
RN Proc. Natl. Acad. Sci. U.S.A. 85:1277-1281(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88118907; PubMed=3430604;
RA Barraclough R., Savin J., Dube S.K., Rudland P.S.;
RT "Molecular cloning and sequence of the gene for p9Ka. A cultured
RT myoepithelial cell protein with strong homology to S-100, a calcium-
RL binding protein.";
RN J. Mol. Biol. 198:13-20(1987).
RN [3]
RP SEQUENCE OF 3-101 FROM N.A.
RX MEDLINE=92158347; PubMed=1741158;
RA de Vouge M.W., Mukherjee B.B.;
RT "Transformation of normal rat kidney cells by v-K-ras enhances
RT expression of transin 2 and an S-100-related calcium-binding
RL protein.";
RN Oncogene 7:109-119(1992).
CC -1- INDUCTION: By nerve growth factor.
CC -1- SIMILARITY: Belongs to the S-100 family.
CC -1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X06916; CAA30014.1; -;
CC EMBL; J03628; AAA42098.1; -;
CC EMBL; X64022; -; NOT ANNOTATED CDS.
CC EMBL; X64023; -; NOT ANNOTATED CDS.
CC PIR; S01759; S01759.
CC HSP; P30801; IAO3.
CC InterPro; IPR001751; CaBP_S100.
CC InterPro; IPR002048; EF-hand.

Thu Sep 9 12:03:20 2004

us-10-067-618-2.rsp

61 DLBNSDQVDFQYAVFLALITIMCNDFQGGSP 94

```

DR Pfam; PF00336; ehand; 1.
DR Pfam; PF01023; S100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProSITE; PS00018; EF_HAND; 1.
DR ProSITE; PS00303; S100_CABP; 1.
KW Calcium-binding.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SQ SEQUENCE 101 AA; 11776 MW; EA0619CEB4F487C1 CRC64;

Query Match          91.6%; Score 488; DB 1; Length 101;
Best Local Similarity 91.1%; Pred. No. 2.8e-39; Indels 0; Gaps 0;
Matches 92; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTFTHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
DQ 1 MARPLEALDVIVSTFTHKYSNGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
OY 61 NLDNRNEVDVDFQYCVFLSCIAMMNCNEFFFGPDQKPRKK 101
DB 61 NLDNRNEVDVDFQYCVFLSCIAMMNCNEFFFGPDQKPRKK 101

RESULT 5
ID S102 BOVIN STANDARD; PRT; 97 AA.
AC P10462;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE S100 calcium-binding protein A2 (S-100L protein).
GN S100A2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP TISSUE=lung;
RX MEDLINE=89139574; PubMed=2521861;
RA Glenney J.R. Jr., Kindy M.S., Zokas L.;
RT "Isolation of a new member of the S100 protein family: amino acid
RT sequence, tissue, and subcellular distribution.";
RL J. Cell Biol. 108:569-578(1989).
CC -!- SUBUNIT: Homomultimeric (Probable).
CC -!- MISCELLANEOUS: This protein binds two calcium ions.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC PIR; A30129; A30129.
DR HSSP; P30801; LA03.
DR InterPro; IPR001751; CaBP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF01023; S100; 1.
DR ProDom; PD003407; CaBP_S100; 1.
DR ProDom; PD000012; EF-hand; 1.
DR ProSITE; PS00018; EF_HAND; 1.
DR ProSITE; PS00303; S100_CABP; 1.
KW Calcium-binding.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 63 74 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
SQ SEQUENCE 97 AA; 10893 MW; 55AC9F60CF9C9E6 CRC64;

Query Match          63.0%; Score 336; DB 1; Length 97;
Best Local Similarity 64.9%; Pred. No. 4.9e-25;
Matches 61; Conservative 17; Mismatches 16; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTFTHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
DQ 1 MSSPLEQALAVVATFTHKYSQGGDKFKLSKGENKELLHKLPSFVGKVDGGLKLMLG 60
OY 61 NLDNRNEVDVDFQYCVFLSCIAMMNCNEFFFGFP 94

```

```

Db RESULT 6
ID S102 HUMAN STANDARD; PRT; 97 AA.
AC P29034; Q00266; Q9BU83;
DT 01-DEC-1992 (Rel. 24, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE S100 calcium-binding protein A2 (S-100L protein) (CANL9).
GN S100A2 OR S100L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN SEQUENCE FROM N.A.
RP MEDLINE=92196147; PubMed=1372446;
RX Lee S.W., Tomasetto C., Swisshelm K., Keyomarsi K., Sager R.;
RT "Down-regulation of a member of the S100 gene family in mammary
RT carcinoma cells and reexpression by azadeoxycytidine treatment.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:2504-2508(1992).
RN SEQUENCE FROM N.A.
RP MEDLINE=98142474; PubMed=9481475;
RX Wikki R., Franz C., Scholl F.A., Heizmann C.W., Schaefer B.M.;
RT "Repression of the candidate tumor suppressor gene S100A2 in breast
RT cancer is mediated by site-specific hypermethylation.";
RL Cell Calcium 22:243-254(1997).
RN SEQUENCE FROM N.A.
RP TISSUE=Ovary;
RX MEDLINE=23288257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max A.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshikiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Rodriguez S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Marra M.A.;
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN SEQUENCE OF 36-39 AND 41-48.
RP TISSUE=Keratinocytes;
RX MEDLINE=93162043; PubMed=1286667;
RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,
RA Vandekerckhove J.;
RT "Microsequences of 145 proteins recorded in the two-dimensional gel
RT protein database of normal human epidermal keratinocytes.";
RL Electrophoresis 13:960-969(1992).
CC -!- FUNCTION: May act as a modulator against excess calcium
CC accumulation in normal human mammary epithelial cells. May also
CC play a role in suppressing tumor cell growth.
CC -!- TISSUE SPECIFICITY: A subset of epithelial cells amongst which
CC normal human mammary epithelial cells and keratinocytes.
CC -!- DEVELOPMENTAL STAGE: Preferentially expressed in normal human
CC mammary epithelial cells as opposed to tumor-derived ones. The
CC level of S100L was shown to correlate inversely with tumor
CC progression.
CC -!- INDUCTION: By growth factors in early G1 phase and probably by

```


"Structural and functional analysis of a growth-regulated gene, the human calyculin.";
J. Biol. Chem. 262:8325-8332(1987).
[4]
SEQUENCE FROM N.A.
RA Wu J., Liu W., Zhou Y., Zhao Z., Peng X., Yuan J., Qiang B.;
RA "Cloning of human calyculin and calyculin binding protein (Cacybp).";
RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A.
RA TISSUE=Brain, and Placenta;
RA MEDLINE=2238257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins B.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Statchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kerteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grinnwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length
RA human and mouse cDNA sequences.";
RA Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[6]
SEQUENCE OF 57-74.
RA MEDLINE=89374276; PubMed=2775283;
RA Gabius H.J., Bardosi A., Gabius S., Hellmann K.P., Karas M.,
RA Kratzin H.;
RA "Identification of a cell cycle-dependent gene product as a sialic
RA acid-binding protein.";
RA Biochem. Biophys. Res. Commun. 163:506-512(1989).
[7]
INTERACTION WITH SUGT1.
RA MEDLINE=22741750; PubMed=12746458;
RA Nowotny M., Spiechowicz M., Jastrzebska B., Filipek A., Kitagawa K.,
RA Kuznicki J.;
RA "Calcium-regulated interaction of Sgt1 with S100A6 (calyculin) and
RA other S100 proteins.";
RA J. Biol. Chem. 278:26923-26928(2003).
CC -|- SUBUNIT: Homodimer, antiparallel. Interacts with SUGT1.
CC stimulated to proliferate. It is inducible by growth factors and
CC overexpressed in acute myeloid leukemias.
CC -|- PTM: The N-terminus is blocked.
CC -|- MISCELLANEOUS: This protein co-purified with the prolactin
CC receptor.
CC -|- SIMILARITY: Belongs to the S-100 family.
CC -|- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC ENBL; M18981; AAA51906.1; -;
CC ENBL; M14300; AAA35886.1; -;
CC ENBL; J02763; AAA51905.1; -;
CC ENBL; AY034480; AAK59702.1; -;
CC ENBL; BC001431; AAH01431.1; -;
CC ENBL; BC009017; AAH09017.1; -;
CC PIR; A28363; BCHUY.

PDB; 1K8U; 10-APR-02.
DR PDB; 1K96; 10-APR-02.
DR PDB; 1K9K; 10-APR-02.
DR Genew; HGNC:10496; S100A6.
DR MIM; 114110; -;
DR GO; GO:0005635; C:nuclear membrane; NAS.
DR GO; GO:0005509; F:calcium ion binding; NAS.
DR GO; GO:0005515; F:protein binding; IPI.
DR GO; GO:0007409; P:axogenesis; NAS.
DR GO; GO:0048146; P:positive regulation of fibroblast proliferation; NAS.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S100; 1.
DR PRODOM; PD003407; CABP_S100; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CABP; 1.
KW Mitogen; Cell cycle; Calcium-binding; Polymorphism; 3D-structure.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT VARIANT 27 27 H -> R (in dbSNP:11974).
FT VARIANT 69 69 N -> S (in dbSNP:1802581).
FT VARIANT 83 83 I -> T (in dbSNP:1802582).
FT VARIANT 83 83 /FTID=VAR_011984.
SQ SEQUENCE 90 AA; 10180 MW; 860CB1416ACBCA1 CRC64;
Query Match 46.5%; Score 248; DB 1; Length 90;
Best Local Similarity 50.0%; Pred. No. 8.2e-17;
Matches 46; Conservative 21; Mismatches 23; Indels 2; Gaps 1;
QY 1 MACPLEKALDVWVSTFHYSGKGDKFKLNSKSELKELTRELPSFLGKRTDEAFQKLS 60
DB 1 MACPLDQALGLLVAIFHYSGREGDKHYLSKELKELT--IGSKLQDAETARLME 58
QY 61 NLDSNRDNEVDPFQYCVFLSCIAMMCMNEFFEG 92
DB 59 DLDKRNKQEVNFQYVTVFLGALALYNEALKG 90
RESULT 9
S106 MOUSE
ID S106 MOUSE STANDARD; PRT; 89 AA.
AC P14069;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Calyculin (prolactin receptor associated protein) (5B10).
GN S100A6 OR CACY.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=93194967; PubMed=8449996;
RA Timmons P.M., Chan C.T.J., Rigby P.W.J., Poirier F.;
RA "The gene encoding the calcium binding protein calyculin is expressed
RA at sites of exocytosis in the mouse.";
RL J. Cell Sci. 104:187-196(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91120152; PubMed=2177633;
RA Guo X., Chambers A.F., Parfett C.L., Waterhouse P., Murphy L.C.,
RA Reid R.E., Craig A.M., Edwards D.R., Denhardt D.T.;
RA "Identification of a serum-inducible messenger RNA (5B10) as the
RA mouse homologue of calyculin: tissue distribution and expression in
RA metastatic, ras-transformed NIH 3T3 cells.";
RL Cell Growth Differ. 1:333-338(1990).
RN [3]
RP SEQUENCE FROM N.A.

RC STRAIN-TUC CFI;
RX MEDLINE=911162002; PubMed=20022257;
RA Wood L., Garter D., Mills M., Hatzembuhler N., Vogeli G.;
RT "Expression of calyculin, a calcium-binding protein, in the
RL keratinous region of growing hair follicles";
RN J. Invest. Dermatol. 96:383-387(1991).
RP SEQUENCE FROM N.A.
RC TISSUE=Breast tumor, and Colon;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Mariani K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.B.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RL human and mouse cDNA sequences";
RN Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RP SEQUENCE OF 56-69.
RX MEDLINE=90088446; PubMed=2597136;
RA Kuznicki J., Filipek A., Hunziker P.E., Huber S., Heizmann C.W.;
RT "Calcium-binding protein from mouse Ehrlich ascites-tumour cells is
RL homologous to human calyculin";
RN Biochem. J. 263:951-956(1989).
RP SEQUENCE OF 6-89, AND INTERACTION WITH OTHER PROTEINS.
RX MEDLINE=91153321; PubMed=1991917;
RA Filipek A., Gerke V., Weber K., Kuznicki J.;
RT "Characterization of the cell-cycle-regulated protein calyculin from
RL Ehrlich ascites tumor cells. Identification of two binding proteins
RT obtained by Ca2(+)-dependent affinity chromatography";
RN Eur. J. Biochem. 195:795-800(1991).
RP PARTIAL SEQUENCE OF 24-33 AND 37-89.
RX MEDLINE=91339739; PubMed=1874170;
RA Thordarson G., Southard J.N., Talamantes F.;
RT "Purification and characterization of mouse decidua calyculin: a
RL novel stimulator of mouse placental lactogen-II secretion";
RN Endocrinology 129:1257-1265(1991).
CC -!- FUNCTION: Seems to interact with annexin II and glyceraldehyde-
CC 3-phosphate dehydrogenase in a calcium-dependent manner.
CC -!- SUBUNIT: Monomer or disulfide-linked dimers. Interacts with SUGT1
CC (By similarity).
CC -!- MISCELLANEOUS: Binds two calcium ions.
CC -!- SIMILARITY: Belongs to the S-100 family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
DR EMBL; M37761; AAA37358.1; -
DR EMBL; X66449; CAA47065.1; -
DR EMBL; X52278; CAA36521.1; -
DR EMBL; BC003832; AAH03832.1; -
DR EMBL; BC010774; AAH10774.1; -

DR PIR; A54314; A54314.
DR HSP; P30801; IA03.
DR MGD; MGI:1339467; S100a6.
DR GO; GO:0005635; C:nuclear membrane; ISS.
DR GO; GO:0005509; F:calcium ion binding; ISS.
DR GO; GO:0005515; P:protein binding; ISS.
DR GO; GO:0007409; P:axogenesis; ISS.
DR GO; GO:0048146; P:positive regulation of fibroblast prolifera. .; ISS.
DR InterPro; IPR001751; CAPP S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CAPP_S100; 1.
DR PROSITE; PS00018; EF_HAND; 1.
DR PROSITE; PS00303; S100_CAPP; 1.
KW Mitogen; Cell cycle; Calcium-binding.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT CONFLICT 51 51 A -> R (IN REF. 2).
SQ SEQUENCE 89 AA; 10051 MW; 2AA1A4018B8CA687 CRC64;
Query Match 46.2%; Score 246; DB 1; Length 89;
Best Local Similarity 52.3%; Pred. No. 1.2e-16;
Matches 46; Conservative 19; Mismatches 21; Indels 2; Gaps 1;
Qy 1 MACPLEKALDVMSTFHYKSGKGFKLKSELKELITRELPSFLGKRTDEAAAFQKLMS 60
Db 1 MACPLDQAILLVAFHFKYSGKGDHLSKKELIKELT--IGSKLODAEIAFLAMD 58
Qy 61 NLDSNRDNEVDFOEYCYVFLSCIAMMCNE 88
Db 59 DLDRNKDQEVNFOEYVAFGLGALAIYNE 86
RESULT 10
S106_RABIT STANDARD; PRT; 90 AA.
ID S106_RABIT
AC P30801;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Calyculin (lung 10 kDa protein).
GN S100a6.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93093152; PubMed=1459239;
RA Ando Y., Watanabe M., Akatsuka H., Tokumitsu H., Hidaka H.;
RT "Site-directed mutation makes rabbit calyculin dimer";
RL FEBS Lett. 314:109-113(1992).
RN [2]
RP SEQUENCE OF 23-31; 36-48 AND 57-89.
RC TISSUE=Lung;
RX MEDLINE=91378440; PubMed=1898017;
RA Tokumitsu H., Kobayashi R., Hidaka H.;
RT "A calcium-binding protein from rabbit lung cytosol identified as the
RL product of growth-regulated gene (2A9) and its binding proteins";
RN Arch. Biochem. Biophys. 288:202-207(1991).
RN [3]
RP ERRATUM.
RX MEDLINE=92061074; PubMed=1952954;
RA Tokumitsu H., Kobayashi R., Hidaka H.;
RL Arch. Biochem. Biophys. 291:401-401(1991).
RN [4]
RP STRUCTURE BY NMR.
RC TISSUE=Lung;
RX MEDLINE=96018617; PubMed=7552751;
RA Potts B.C.M., Smith J., Akke M., Macke T.J., Okazaki K., Hikada H.,
RA Case D.A., Chazin W.J.;
RT "The structure of calyculin reveals a novel homodimeric fold for S100

```
RT Ca(2+)-binding proteins." ;
RL Nat. Struct. Biol. 2:790-796(1995).
RN [5]
RP STRUCTURE BY NMR.
RX MEDLINE=98179938; PubMed=9519412;
RA Sastry M., Ketchum R.R., Crescenzi O., Weber C., Lubinski M.J.,
RA Hikada H., Chazin W.J.;
RT "The three-dimensional structure of Ca(2+)-bound calyculin:
RT implications for Ca(2+)-signal transduction by S100 proteins." ;
RL Structure 6:223-231(1998).
CC -1- SUBUNIT: Homodimer, antiparallel. Interacts with SUGT1.
CC -1- SIMILARITY: Belongs to the S-100 family.
CC -1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; D10885; BAA01707.1; -.
DR PIR; S27011; S27011.
DR PDB; ICNP; 14-OCT-96.
DR PDB; 1A03; 02-MAR-99.
DR PDB; 1JWD; 27-MAR-02.
DR PDB; 2CNP; 22-JUL-99.
DR GO; GO:0005635; C:nuclear membrane; ISS.
DR GO; GO:0005503; F:calcium ion binding; ISS.
DR GO; GO:0005515; F:protein binding; ISS.
DR GO; GO:0007403; P:axogenesis; ISS.
DR GO; GO:0048146; P:positive regulation of fibroblast proliferation. . ; ISS.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR Pfam; PF01023; S_100; 1.
DR ProDom; PD003407; CABP_S100; 1.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100 CABP; 1.
KW Mitogen; Cell cycle; Calcium-binding; 3D-structure.
FT CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
FT CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
FT HELIX 4 19
FT TURN 20 23
FT STRAND 29 30
FT HELIX 31 42
FT TURN 43 45
FT TURN 47 47
FT HELIX 48 60
FT STRAND 67 68
FT HELIX 70 83
FT TURN 84 84
FT HELIX 85 90
SQ SEQUENCE 90 AA; 10154 MW; 9B23BC724B9E771F CRC64;
Query Match 45.8%; Score 244; DB 1; Length 90;
Best Local Similarity 51.1%; Pred. No. 1.9e-16;
Matches 47; Conservative 18; Mismatches 25; Indels 2; Gaps 1;
QY 1 MACPLEKALDVMTSTPHKSGKGDGFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMNS 60
Db 1 MASPLDQALGLLIGIPHKSGKGDGKHTLSKKELKELIQKELT--IGSKLQDAEIVKLM 58
QY 61 NLDNRDNEVDFOEYCVFVLSCIAMMCNEFFEG 92
Db 59 DLDNRKQDQVNFQYITFTLGALAMINYEALKG 90
RESULT 11
S10A BOVIN STANDARD; PRT; 93 AA.
AC P02639;
```

```
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE S-100 protein, alpha chain.
GN S100A1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86248083; PubMed=3755105;
RA Kuwano R., Maeda T., Usui H., Araki K., Yamakuni T., Ohshima Y.,
RA Kurihara T., Takahashi Y.;
RT "Molecular cloning of cDNA of S100 alpha subunit mRNA." ;
RL FEBS Lett. 202:97-101(1986).
RN [2]
RP SEQUENCE.
RX MEDLINE=81236562; PubMed=7250124;
RA Isobe T., Okuyama T.;
RT "The amino-acid sequence of the alpha subunit in bovine brain S-100a
RT protein." ;
RL Eur. J. Biochem. 116:79-86(1981).
RN [3]
RP METAL ION-BINDING PROPERTIES.
RX MEDLINE=84000339; PubMed=6615778;
RA Baudier J., Gerard D.;
RT "Ions binding to S100 proteins: structural changes induced by calcium
RT and zinc on S100a and S100b proteins." ;
RL Biochemistry 22:3360-3369(1983).
CC -1- FUNCTION: Weakly binds calcium but binds zinc very tightly-
CC distinct binding sites with different affinities exist for both
CC ions on each monomer. Physiological concentrations of potassium
CC ion antagonize the binding of both divalent cations, especially
CC affecting high-affinity calcium-binding sites.
CC -1- SUBUNIT: Dimer of either two alpha chains, or two beta chains, or
CC one alpha and one beta chain.
CC -1- TISSUE SPECIFICITY: Although predominant among the water-soluble
CC brain proteins, S-100 is also found in a variety of other tissues.
CC -1- SIMILARITY: Belongs to the S-100 family.
CC -1- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
DR PIR; A24156; BCBO1A.
DR HSP; P04631; 1B4C.
DR InterPro; IPR001751; CABP_S100.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; ehand; 1.
DR ProDom; PD003407; CABP_S100; 1.
DR PROSITE; PS00018; EF HAND; 1.
DR PROSITE; PS00303; S100 CABP; 1.
KW Calcium-binding; Zinc; Metal-binding.
FT INIT_MET 0 0
FT MOD_RES 1 1 BLOCKED.
FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY).
FT CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY).
FT CONFLICT 64 64 N -> D (IN REF. 2).
FT SEQUENCE 93 AA; 10387 MW; 0457D0F44819B89B CRC64;
Query Match 45.4%; Score 242; DB 1; Length 93;
Best Local Similarity 52.3%; Pred. No. 3.1e-16;
Matches 45; Conservative 14; Mismatches 27; Indels 0; Gaps 0;
QY 5 LEKALDVMTSTPHKSGKGDGFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMNS 64
Db 4 LETAMETLNVFHAHSGKGDGKHTLSKKELKELIQKELT--IGSKLQDAEIVKLM 63
QY 65 NRDNRDNEVDFOEYCVFVLSCIAMMCNEFF 90
Db 64 NGDGEVDFOEYVVLVAALTVAACNFF 89
RESULT 12
```


FT INIT_MET 0 0 BY SIMILARITY.
 FT CA_BIND 19 32 EF-HAND 1 (LOW AFFINITY).
 FT CA_BIND 62 73 EF-HAND 2 (HIGH AFFINITY).
 SQ SEQUENCE 93 AA; 10415 MW; 0457C356C222B89B CRC64;
 Query Match 45.2%; Score 241; DB 1; Length 93;
 Best Local Similarity 52.3%; Pred. No. 3.8e-16;
 Matches 45; Conservative 14; Mismatches 27; Indels 0; Gaps 0;
 5 LEKALDVMVSTFKYSGKDGKFLNKSSELKELTRELPSFLGKRTDAAAFQKLMSNLS 64
 4 LETAMETLVNFHAHSGKDGKFLNKSSELKELTRELPSFLGKRTDAAAFQKLMSNLS 63
 65 NRDNVDQFQYCVFLSCIAMMNCNFF 90
 64 NGDGEVDQFQYCVFLSCIAMMNCNFF 89
 RESULT 14
 S106_CHICK
 ID S106_CHICK STANDARD; PRT; 92 AA.
 AC Q98953;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Calcyclin.
 GN S100A6.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Gizzard;
 RX MEDLINE=98262396; PubMed=9599662;
 RA Allen B.G., Andrea J.E., Sutherland C., Schonekess B.O., Walsh M.P.;
 RT "Molecular cloning of chicken calcyclin (S100A6) and identification
 of putative isoforms.";
 RL Biochem. Cell Biol. 75:733-738(1997).
 CC -!- SIMILARITY: Belongs to the S-100 family.
 CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U76365; AAB18788.1; -.
 CC HSP; P30801; 1A03.
 CC GO; GO:0005635; C:nuclear membrane; ISS.
 CC GO; GO:0005509; F:calcium ion binding; ISS.
 CC GO; GO:0005515; F:protein binding; ISS.
 CC GO; GO:0007409; P:axogenesis; ISS.
 CC GO; GO:0048146; P:positive regulation of fibroblast proliferation. . .; ISS.
 CC InterPro; IPR001751; CaBP S100.
 CC InterPro; IPR002048; EF-hand.
 CC Pfam; PF00036; ehand; 1.
 CC Pfam; PF01023; S_100; 1.
 CC ProDom; PD003407; CaBP_S100; 1.
 CC PROSITE; PS00018; EF HAND; 1.
 CC PROSITE; PS00303; S100 CBP; 1.
 CC Mitogen; Cell cycle; Calcium-binding.
 CC CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY) (POTENTIAL).
 FT CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY) (POTENTIAL).
 SQ SEQUENCE 92 AA; 10276 MW; 56B3D65BA7BF7A4 CRC64;
 Query Match 44.8%; Score 239; DB 1; Length 92;
 Best Local Similarity 53.4%; Pred. No. 5.8e-16;
 Matches 47; Conservative 17; Mismatches 22; Indels 2; Gaps 1;
 1 MACPLEKALDVMVSTFKYSGKDGKFLNKSSELKELTRELPSFLGKRTDAAAFQKLMS 60
 1 MAEPDQAIGLVATFVKYSGKDGKFLNKSSELKELTRELPSFLGKRTDAAAFQKLMS 58
 61 NLDNRNVDQFQYCVFLSCIAMMNCNE 88
 59 DLDNRNVDQFQYCVFLSCIAMMNCNE 86
 RESULT 15
 S105_MOUSE
 ID S105_MOUSE STANDARD; PRT; 93 AA.
 AC O88945; P82540;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE S100 calcium-binding protein A5 (S-100D protein).
 GN S100A5 OR S100D.
 OS Mus musculus (Mouse), and
 OC Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090, 10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=Mouse;
 RX MEDLINE=99117144; PubMed=9920416;
 RA Ridinger K., Ilg E.C., Niggli F.K., Heizmann C.W., Schaefer B.W.;
 RT "Clustered organization of S100 genes in human and mouse.";
 RL Biochim. Biophys. Acta 1448:254-263(1998).
 RN [2]
 RP SEQUENCE, FUNCTION, SUBUNIT, AND MASS SPECTROMETRY.
 RC SPECIES= Rat; STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=20459141; PubMed=10882717;
 RA Schaefer B.W., Fritschy J.-M., Murrmann P., Troxler H., Durussel I.;
 RT "Brain S100A5 is a novel calcium-, zinc-, and copper ion-binding
 protein of the EF-hand superfamily.";
 RL J. Biol. Chem. 275:30623-30630(2000).
 CC -!- FUNCTION: Binds calcium, zinc and copper. Binds 2 calcium and
 copper ions per molecule and 1 zinc ion per molecule.
 CC -!- SUBUNIT: Homodimer.
 CC -!- MASS SPECTROMETRY: MW=10878; METHOD=Electrospray.
 CC -!- SIMILARITY: Belongs to the S-100 family.
 CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF087469; AAC64106.1; -.
 CC HSP; P30801; 1A03.
 CC MG1; MG1:1338915; S100a5.
 CC InterPro; IPR001751; CaBP S100.
 CC InterPro; IPR002048; EF-hand.
 CC Pfam; PF00036; ehand; 1.
 CC Pfam; PF01023; S_100; 1.
 CC ProDom; PD003407; CaBP_S100; 1.
 CC PROSITE; PS00018; EF HAND; 1.
 CC PROSITE; PS00303; S100 CBP; 1.
 CC Mitogen; Cell cycle; Calcium-binding.
 CC CA_BIND 20 33 EF-HAND 1 (LOW AFFINITY).
 FT CA_BIND 61 72 EF-HAND 2 (HIGH AFFINITY).
 SQ SEQUENCE 93 AA; 10812 MW; 9A922E1898D202A4 CRC64;
 Query Match 44.8%; Score 239; DB 1; Length 93;
 Best Local Similarity 52.2%; Pred. No. 5.9e-16;

Matches 47; Conservative 15; Mismatches 26; Indels 2; Gaps 1;
QY 1 MACPLEKALDVWVSTFHKYSRGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
Db ||||| ||:|||||:| |::|||: || |:: ||
1 METPLEKALTWVTFHFKYSRGEGSKLTLSRKELKELIKTELS--LAEXKXESSIDNLMK 58
QY 61 NLDSNRDNEVDFOEYCVFLSCIAMMCNEFF 90
Db :|| | |::||| |||: | |::||

Search completed: September 9, 2004, 10:53:48
Job time : 29.671 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:53:18 ; Search time 135,535 Seconds
(without alignments)
238.975 Million cell updates/sec

Title: US-10-067-618-2

Perfect score: 533

Sequence: 1 MACPLEKALDVMVSTPHKYS.....IAMCNNEFFGPDQPRKK 101

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1335176 seqs, 320689617 residues

Total number of hits satisfying chosen parameters: 1335176

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published Applications_AA:*

```

1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	533	100.0	101	9	US-09-393-433-1
2	533	100.0	101	9	US-09-781-509-1
3	533	100.0	101	12	US-10-087-192-1158
4	533	100.0	101	13	US-10-067-618-2
5	533	100.0	101	13	US-10-135-152-2
6	533	100.0	101	14	US-10-269-643-1
7	496	93.1	101	9	US-09-393-433-2
8	496	93.1	101	9	US-09-781-509-2
9	496	93.1	101	14	US-10-269-643-2
10	496	93.1	119	12	US-10-087-192-1155
11	323	60.6	97	14	US-10-097-340-274
12	323	60.6	97	14	US-10-171-311-206
13	323	60.6	97	15	US-10-236-031B-40
14	248	46.5	90	9	US-09-738-973-200
15	248	46.5	90	9	US-09-974-298-20

16	248	46.5	90	9	US-09-854-133-200	Sequence 200, Appl
17	248	46.5	90	14	US-10-097-340-276	Sequence 276, Appl
18	248	46.5	90	14	US-10-144-649A-200	Sequence 200, Appl
19	246	46.2	105	14	US-10-106-698-4570	Sequence 4570, Appl
20	246	46.2	105	15	US-10-264-049-2219	Sequence 2219, Appl
21	244	45.8	90	16	US-10-408-765A-2429	Sequence 2429, Appl
22	244	45.8	94	14	US-10-097-340-270	Sequence 270, Appl
23	231	43.3	97	16	US-10-363-829-412	Sequence 412, Appl
24	229	43.0	92	10	US-09-492-026-5	Sequence 5, Appl
25	229	43.0	92	10	US-09-919-039-184	Sequence 184, Appl
26	217	40.7	92	12	US-10-336-603A-102	Sequence 102, Appl
27	216	40.5	73	12	US-10-424-599-233903	Sequence 233903, Appl
28	214	40.2	93	15	US-10-094-886-58	Sequence 102, Appl
29	200	37.5	95	9	US-09-919-172-102	Sequence 58, Appl
30	200	37.5	95	9	US-09-981-353-98	Sequence 102, Appl
31	200	37.5	113	12	US-10-276-774-2377	Sequence 98, Appl
32	183	34.3	91	14	US-10-106-698-6907	Sequence 2377, Appl
33	174.5	32.7	97	10	US-09-877-843-28	Sequence 6907, Appl
34	171.5	32.2	97	10	US-09-877-843-29	Sequence 28, Appl
35	170.5	32.0	97	10	US-09-877-843-26	Sequence 29, Appl
36	170.5	32.0	97	11	US-09-997-003-32	Sequence 26, Appl
37	170.5	32.0	97	14	US-10-304-287-3	Sequence 32, Appl
38	170.5	32.0	97	16	US-10-735-577-3	Sequence 3, Appl
39	170.5	32.0	118	15	US-10-264-049-3289	Sequence 3, Appl
40	170.5	32.0	134	10	US-09-992-600A-2	Sequence 3289, Appl
41	170.5	32.0	134	10	US-09-924-340-2	Sequence 2, Appl
42	170.5	32.0	134	10	US-09-992-095B-2	Sequence 2, Appl
43	170.5	32.0	134	10	US-09-999-570-2	Sequence 2, Appl
44	170.5	32.0	134	14	US-10-000-489-2	Sequence 2, Appl
45	170.5	32.0	134	14	US-10-000-986-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-09-393-433-1
; Sequence 1, Application US/09393433
; Patent No. US20010011126A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-393-433-1

Query Match 100.0%; Score 533; DB 9; Length 101;
Best Local Similarity 100.0%; Pred. No. 6e-54;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MACPLEKALDVMVSTPHKYSKGEGDKFLNKSELKELLTRELPSFLGKRTDAAAFQKLMS 60
Db 1 MACPLEKALDVMVSTPHKYSKGEGDKFLNKSELKELLTRELPSFLGKRTDAAAFQKLMS 60
QY 61 NLDNRDNEVDFOEYCVFLSCIAMCNNEFFGPDQPRKK 101
Db 61 NLDNRDNEVDFOEYCVFLSCIAMCNNEFFGPDQPRKK 101

RESULT 2
US-09-781-509-1
; Sequence 1, Application US/09781509
; Patent No. US20020095010A1
; GENERAL INFORMATION:

```

; APPLICANT: Lukanidin, Eugene
; APPLICANT: Bock, Elisabeth M.
; APPLICANT: Berezen, Vladimir
; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS
; FILE REFERENCE: RCT
; CURRENT APPLICATION NUMBER: US/09/781,509
; PRIOR FILING DATE: 2001-02-12
; PRIOR APPLICATION NUMBER: 09/393,433
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-781-509-1

Query Match      100.0%; Score 533; DB 9; Length 101;
Best Local Similarity 100.0%; Pred. No. 6e-54;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MACPLEKALDVWVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db      1 MACPLEKALDVWVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60

Qy      61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101
Db      61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101

RESULT 3
US-10-087-192-1158
; Sequence 1158, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1158
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-1158

Query Match      100.0%; Score 533; DB 12; Length 101;
Best Local Similarity 100.0%; Pred. No. 6e-54;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MACPLEKALDVWVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db      1 MACPLEKALDVWVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60

Qy      61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101
Db      61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101

RESULT 4
US-10-067-618-2
; Sequence 2, Application US/10067618
; Publication No. US20020169110A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene

```

```

; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 7879ZYA111-Z
; CURRENT APPLICATION NUMBER: US/10/067,618
; CURRENT FILING DATE: 2002-02-05
; PRIOR APPLICATION NUMBER: 09/298,625
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-067-618-2

Query Match      100.0%; Score 533; DB 13; Length 101;
Best Local Similarity 100.0%; Pred. No. 6e-54;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MACPLEKALDVWVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db      1 MACPLEKALDVWVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60

Qy      61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101
Db      61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101

RESULT 5
US-10-135-152-2
; Sequence 2, Application US/10135152
; Publication No. US20020172680A1
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 7879ZYA111-Z
; CURRENT APPLICATION NUMBER: US/10/135,152
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US/09/298,625
; PRIOR FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-135-152-2

Query Match      100.0%; Score 533; DB 13; Length 101;
Best Local Similarity 100.0%; Pred. No. 6e-54;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MACPLEKALDVWVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db      1 MACPLEKALDVWVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60

Qy      61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101

```


Db 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101

RESULT 6

US-10-269-643-1

; Sequence 1, Application US/10269643

; Publication No. US20030100503A1

; GENERAL INFORMATION:

; APPLICANT: Lukanidin, Eugene

; APPLICANT: Bock, Elisabeth M.

; APPLICANT: Berezen, Vladimir

; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS

; FILE REFERENCE: RCT

; CURRENT APPLICATION NUMBER: US/10/269,643

; CURRENT FILING DATE: 2002-10-11

; PRIOR APPLICATION NUMBER: US/09/781,509

; PRIOR FILING DATE: 2001-02-12

; PRIOR APPLICATION NUMBER: 09/393,433

; PRIOR FILING DATE: 1999-09-10

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 1

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-269-643-1

Query Match 100.0%; Score 533; DB 14; Length 101;

Best Local Similarity 100.0%; Pred. No. 1.2e-54;

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTSTFHKYSKGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60

Db 1 MACPLEKALDVMTSTFHKYSKGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60

QY 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101

Db 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101

RESULT 7

US-09-393-433-2

; Sequence 2, Application US/09393433

; Patent No. US20010011126A1

; GENERAL INFORMATION:

; APPLICANT: Lukanidin, Eugene

; APPLICANT: Bock, Elisabeth M.

; APPLICANT: Berezen, Vladimir

; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS

; FILE REFERENCE: RCT

; CURRENT APPLICATION NUMBER: US/09/393,433

; CURRENT FILING DATE: 1999-09-10

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-393-433-2

Query Match 93.1%; Score 496; DB 9; Length 101;

Best Local Similarity 93.1%; Pred. No. 1.2e-49;

Matches 94; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTSTFHKYSKGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60

Db 1 MACPLEKALDVMTSTFHKYSKGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60

QY 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101

Db 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101

RESULT 8

US-09-781-509-2

; Sequence 2, Application US/09781509

; Patent No. US2002009010A1

; GENERAL INFORMATION:

; APPLICANT: Lukanidin, Eugene

; APPLICANT: Bock, Elisabeth M.

; APPLICANT: Berezen, Vladimir

; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS

; FILE REFERENCE: RCT

; CURRENT APPLICATION NUMBER: US/09/781,509

; CURRENT FILING DATE: 2001-02-12

; PRIOR APPLICATION NUMBER: 09/393,433

; PRIOR FILING DATE: 1999-09-10

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-781-509-2

Query Match 93.1%; Score 496; DB 9; Length 101;

Best Local Similarity 93.1%; Pred. No. 1.2e-49;

Matches 94; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTSTFHKYSKGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60

Db 1 MACPLEKALDVMTSTFHKYSKGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60

QY 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101

Db 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101

RESULT 9

US-10-269-643-2

; Sequence 2, Application US/10269643

; Publication No. US20030100503A1

; GENERAL INFORMATION:

; APPLICANT: Lukanidin, Eugene

; APPLICANT: Bock, Elisabeth M.

; APPLICANT: Berezen, Vladimir

; TITLE OF INVENTION: NEUROGENIC COMPOSITIONS AND METHODS

; FILE REFERENCE: RCT

; CURRENT APPLICATION NUMBER: US/10/269,643

; CURRENT FILING DATE: 2002-10-11

; PRIOR APPLICATION NUMBER: US/09/781,509

; PRIOR FILING DATE: 2001-02-12

; PRIOR APPLICATION NUMBER: 09/393,433

; PRIOR FILING DATE: 1999-09-10

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 2

; LENGTH: 101

; TYPE: PRT

; ORGANISM: Mus musculus

US-10-269-643-2

Query Match 93.1%; Score 496; DB 14; Length 101;

Best Local Similarity 93.1%; Pred. No. 1.2e-49;

Matches 94; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTSTFHKYSKGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60

Db 1 MACPLEKALDVMTSTFHKYSKGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60

QY 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101

Db 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101

RESULT 10

```
US-10-087-192-1155
; Sequence 1155, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-1155

Query Match          93.1%; Score 496; DB 12; Length 119;
Best Local Similarity 93.1%; Pred. No. 1.5e-49;
Matches 94; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MACPLEKALDVNVSTFHKYSKGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
Db 19 MARPLEALDVIVSTFHKYSKGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKMS 78

Qy 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDK 101
Db 79 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGCPDKPRK 119

RESULT 11
US-10-097-340-274
; Sequence 274, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATYAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102

US-10-097-340-274
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 274
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-274

Query Match          60.6%; Score 323; DB 14; Length 97;
Best Local Similarity 61.5%; Pred. No. 1.4e-29;
Matches 59; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MACPLEKALDVNVSTFHKYSKGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
Db 1 MCSLEQALAVLVITTFHKYSCQEGDKFKLSKGEMKELLHKLPSFVGKVEDEGLKMLG 60

Qy 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDK 96
Db 61 SLDENSQQVDFQEVAVFLALITVMCNDFFQGCPCR 96

RESULT 12
US-10-171-311-206
; Sequence 206, Application US/10171311
; Publication No. US20030087270A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Chen, Yan
; APPLICANT: Zhao, Xumei
; APPLICANT: Monahan, John
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Glatt, Karen
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Hoersh, Sebastian
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
; TITLE OF INVENTION: OF CERVICAL CANCER
; FILE REFERENCE: MRI-035
; CURRENT APPLICATION NUMBER: US/10/171,311
; CURRENT FILING DATE: 2002-06-12
; PRIOR APPLICATION NUMBER: US 60/298,159
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,155
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/335,936
; PRIOR FILING DATE: 2001-11-14
; NUMBER OF SEQ ID NOS: 238
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 206
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-311-206

Query Match          60.6%; Score 323; DB 14; Length 97;
Best Local Similarity 61.5%; Pred. No. 1.4e-29;
Matches 59; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MACPLEKALDVNVSTFHKYSKGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
Db 1 MCSLEQALAVLVITTFHKYSCQEGDKFKLSKGEMKELLHKLPSFVGKVEDEGLKMLG 60

Qy 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDK 96
Db 61 SLDENSQQVDFQEVAVFLALITVMCNDFFQGCPCR 96

RESULT 13
US-10-236-031B-40
; Sequence 40, Application US/10236031B
```

```
US-10-087-192-1155
; Sequence 1155, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1155
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-087-192-1155

Query Match          93.1%; Score 496; DB 12; Length 119;
Best Local Similarity 93.1%; Pred. No. 1.5e-49;
Matches 94; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MACPLEKALDVNVSTFHKYSKGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
Db 19 MARPLEALDVIVSTFHKYSKGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKMS 78

Qy 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDK 101
Db 79 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGCPDKPRK 119

RESULT 11
US-10-097-340-274
; Sequence 274, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATYAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
```

```

; Publication No. US20030219760A1
; GENERAL INFORMATION:
; APPLICANT: Gordon, Gavin J.
; APPLICANT: Jensen, Roderick V.
; APPLICANT: Gullans, Steven R.
; APPLICANT: Bueno, Raphael
; TITLE OF INVENTION: Diagnostic and Prognostic Tests
; FILE REFERENCE: B00801/70265 (JRV/JAV)
; CURRENT APPLICATION NUMBER: US/10/236,031B
; CURRENT FILING DATE: 2002-09-05
; PRIOR APPLICATION NUMBER: US 60/317,389
; PRIOR FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US 60/407,431
; PRIOR FILING DATE: 2002-08-30
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40
; LENGTH: 97
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-236-031B-40

Query Match      60.6%; Score 323; DB 15; Length 97;
Best Local Similarity 61.5%; Pred. No. 1.4e-29;
Matches 59; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

Qy 1 MACPLEKALDVMVSTFHKYSKGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
Db 1 MCSLEQALAVVTFHFKYSQCGDKFKLSKGEMKELHKLPSFGKGVDEGLKLMG 60

Qy 61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDK 96
Db 61 SLDNSNQQVDFQEYAVFLALITWNCNDFQGCEDR 96

```

```

RESULT 14
US-09-738-973-200
; Sequence 200, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Raedoh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliot, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 200
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-738-973-200

```

```

Query Match      46.5%; Score 248; DB 9; Length 90;
Best Local Similarity 50.0%; Pred. No. 6.3e-21;
Matches 46; Conservative 21; Mismatches 23; Indels 2; Gaps 1;

Qy 1 MACPLEKALDVMVSTFHKYSKGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
Db 1 MACPLDQALIGLLVAIFHKYSGREGDKHTLSKKELKELIQKELT--IGSKLQDAEIAFLME 58

```

```

Qy 61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEG 92
Db 59 DLDNRKQDEVNFQBYVTFGLGALALIYNEALKG 90

RESULT 15
US-09-974-298-20
; Sequence 20, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 2748261CD1
US-09-974-298-20

Query Match      46.5%; Score 248; DB 9; Length 90;
Best Local Similarity 50.0%; Pred. No. 6.3e-21;
Matches 46; Conservative 21; Mismatches 23; Indels 2; Gaps 1;

Qy 1 MACPLEKALDVMVSTFHKYSKGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
Db 1 MACPLDQALIGLLVAIFHKYSGREGDKHTLSKKELKELIQKELT--IGSKLQDAEIAFLME 58

Qy 61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEG 92
Db 59 DLDNRKQDEVNFQBYVTFGLGALALIYNEALKG 90

```

Search completed: September 9, 2004, 11:04:43
Job time : 136.535 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:47:43 ; Search time 56.6903 Seconds
(without alignments)
91.977 Million cell updates/sec

Title: US-10-067-618-2

Perfect score: 533

Sequence: 1 MACPLEKALDVMVSTFKYS.....IAMMCNEFFBGFDPKQPRKK 101

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCITUS COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	533	100.0	101	1	US-08-190-560-2
2	533	100.0	101	1	US-08-469-277-2
3	533	100.0	101	2	US-08-468-946-2
4	533	100.0	101	2	US-08-468-942-2
5	533	100.0	101	4	US-09-298-625-2
6	328	61.5	97	1	US-07-662-198B-2
7	248	46.5	90	4	US-09-370-838-200
8	243	45.6	89	1	US-07-987-272A-10
9	240	45.0	92	2	US-09-621-976-7537
10	231	43.3	92	2	US-09-051-589-1
11	229	43.0	91	1	US-07-987-272A-11
12	229	43.0	92	2	US-08-918-727-5
13	229	43.0	92	3	US-09-205-680A-5
14	211	39.6	88	4	US-09-621-976-7524
15	211	39.6	88	4	US-09-621-976-7526
16	210	39.4	88	4	US-09-621-976-7528
17	200	37.5	95	4	US-09-919-172-102
18	200	37.5	95	4	US-09-976-594-467
19	168.5	31.6	95	1	US-07-987-272A-9
20	165.5	31.1	95	4	US-09-399-913-65
21	165.5	31.1	113	3	US-08-918-727-7
22	165.5	31.1	113	3	US-09-205-680A-7
23	159.5	29.9	105	2	US-08-918-727-6
24	159.5	29.9	105	3	US-09-205-680A-6
25	157.5	29.5	92	2	US-08-568-310D-20
26	157.5	29.5	92	4	US-09-270-455-20
27	156	29.3	91	3	US-08-794-000-2

28 153 28.7 90 4 US-09-263-312-3 Sequence 3, Appli
29 153 28.7 90 4 US-09-826-589-3 Sequence 3, Appli
30 153 28.7 90 4 US-09-826-589-4 Sequence 4, Appli
31 153 28.7 92 2 US-08-568-310D-19 Sequence 19, Appli
32 153 28.7 92 4 US-09-270-455-19 Sequence 19, Appli
33 150.5 28.2 114 1 US-08-385-241-3 Sequence 3, Appli
34 150.5 28.2 114 4 US-09-214-272-4 Sequence 4, Appli
35 149 28.0 109 1 US-07-987-272A-8 Sequence 8, Appli
36 147 27.6 64 4 US-09-621-976-4767 Sequence 4767, Ap
37 141 26.5 44 1 US-08-056-200-100 Sequence 100, App
38 141 26.5 44 2 US-08-800-644-100 Sequence 100, App
39 137 25.7 45 1 US-08-056-200-97 Sequence 97, Appl
40 137 25.7 45 2 US-08-800-644-97 Sequence 97, Appl
41 135.5 25.4 103 2 US-08-918-727-1 Sequence 1, Appli
42 135.5 25.4 103 3 US-09-205-680A-1 Sequence 1, Appli
43 135.5 25.4 103 4 US-09-010-147B-2 Sequence 2, Appli
44 134.5 25.2 98 2 US-08-918-727-3 Sequence 3, Appli
45 134.5 25.2 98 3 US-09-205-680A-3 Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-190-560-2
; Sequence 2, Application US/08190560
; Patent No. 5798257
; GENERAL INFORMATION:
; APPLICANT: Zain, Sayeeda
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY
; TITLE OF INVENTION: THE MTS-1 GENE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,560
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 7879ZY
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-190-560-2

Query Match 100.0%; Score 533; DB 1; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.2e-61;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMVSTFKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAFKLMS 60
DB 1 MACPLEKALDVMVSTFKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAFKLMS 60

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/468,942
;; FILING DATE: 06-JUN-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 08/190,560
;; FILING DATE: 31-JAN-1994
;; ATTORNEY/AGENT INFORMATION:
;; NAME: DiGiglio, Frank S.
;; REGISTRATION NUMBER: 31,346
;; REFERENCE/DOCKET NUMBER: 78792Y
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (516) 742-4343
;; TELEFAX: (516) 742-4366
;; TELEX: 230 901 SANS UR
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 101 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; US-08-468-942-2

Query Match 100.0%; Score 533; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.2e-61;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTSTFKYSGKGDGKFLNKSSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db |||||
QY 1 MACPLEKALDVMTSTFKYSGKGDGKFLNKSSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db |||||

QY 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGPDKQPRKK 101
Db |||||

RESULT 5
US-09-298-625-2
; Sequence 2, Application US/09298625
; Patent No. 6638504
; GENERAL INFORMATION:
; APPLICANT: Lukanidin, Eugene
; TITLE OF INVENTION: DIAGNOSIS OF METASTATIC CANCER BY THE MTS-1 GENE
; FILE REFERENCE: 78792YATII-Z
; CURRENT APPLICATION NUMBER: US/09/298,625
; CURRENT FILING DATE: 1999-04-23
; PRIOR APPLICATION NUMBER: 08/468,942
; PRIOR FILING DATE: 1995-06-06
; PRIOR APPLICATION NUMBER: 08/190,560
; PRIOR FILING DATE: 1994-01-31
; PRIOR APPLICATION NUMBER: 07/981,455
; PRIOR FILING DATE: 1992-11-25
; PRIOR APPLICATION NUMBER: 07/550,600
; PRIOR FILING DATE: 1990-07-09
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-298-625-2

Query Match 100.0%; Score 533; DB 4; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.2e-61;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTSTFKYSGKGDGKFLNKSSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db |||||
QY 1 MACPLEKALDVMTSTFKYSGKGDGKFLNKSSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db |||||

QY 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGPDKQPRKK 101
Db |||||

Db 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGPDKQPRKK 101

RESULT 6
US-07-662-198B-2
; Sequence 2, Application US/07662198B
; Patent No. 5262528
; GENERAL INFORMATION:
; APPLICANT: Sager, Ruth
; APPLICANT: Lee, Sam W.
; APPLICANT: Tomasetto, Catherine
; TITLE OF INVENTION: CANCER DIAGNOSIS AND THERAPY
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/662,198B
; FILING DATE: 19910228
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/049001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 97
; TYPE: AMINO ACID
; STRANDEDNESS:
; TOPOLOGY: linear
; US-07-662-198B-2

Query Match 61.5%; Score 328; DB 1; Length 97;
Best Local Similarity 62.5%; Pred. No. 5.2e-35;
Matches 60; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTSTFKYSGKGDGKFLNKSSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db |||||
QY 1 MACPLEKALDVMTSTFKYSGKGDGKFLNKSSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db |||||

QY 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGPDK 96
Db |||||

Db 61 NLDNSDQQVDFQYAVFLALITVMCNDFFQGCPCR 96

RESULT 7
US-09-370-838-200
; Sequence 200, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadoh
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; TITLE OF INVENTION: LUNG CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.475C1

```

; CURRENT APPLICATION NUMBER: US/09/370.838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285.323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 200
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-370-838-200

```

```

Query Match 46.5%; Score 248; DB 4; Length 90;
Best Local Similarity 50.0%; Pred. No. 1.2e-24;
Matches 46; Conservative 21; Mismatches 23; Indels 2; Gaps 1;

QY 1 MACPLEKALDVMTSTFKYSGEGDKFKLNSKELKLTRELPSFLGKRTDEAAFOKLMS 60
DB 1 MACPLDQAIGLLVAIFHKYSGREGDKHTLSKKELKELIKELT--IGSKLQDAEIARLME 58

QY 61 NLDNRDNEVDFOEYCVFLSCIAMMNCNEFFEG 92
DB 59 LDRNKDQEVNFQEVYVTFGLALIIYNEALKG 90

```

```

RESULT 8
US-07-987-272A-10
; Sequence 10, Application US/07987272A
; Patent No. 5731166
; GENERAL INFORMATION:
; APPLICANT: Geczy, C., Simpson, R. J. and Lackmann, M
; TITLE OF INVENTION: No. 5731166el Chemotactic Factor
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cushman Darby & Cushman
; STREET: 1100 New York Avenue, N. W., Ninth Floor, East Tower
; CITY: Washington
; STATE: D. C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/987,272A
; FILING DATE: 05-MAR-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK 2127
; FILING DATE: 05-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PK 4463
; FILING DATE: 05-SEP-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Brinkman, David W
; REGISTRATION NUMBER: 20,817
; REFERENCE/DOCKET NUMBER: DWB/1925/200259
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861 3000
; TELEFAX: 202-822 0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 89 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-987-272A-10

```

Query Match 45.6%; Score 243; DB 1; Length 89;

```

Best Local Similarity 49.5%; Pred. No. 5.3e-24;
Matches 45; Conservative 21; Mismatches 23; Indels 2; Gaps 1;

QY 2 ACPLEKALDVMTSTFKYSGEGDKFKLNSKELKLTRELPSFLGKRTDEAAFOKLMS 61
DB 1 ACPLDQAIGLLVAIFHKYSGREGDKHTLSKKELKELIKELT--IGSKLQDAEIARLME 58

QY 62 LDRNRDNEVDFOEYCVFLSCIAMMNCNEFFEG 92
DB 59 LDRNKDQEVNFQEVYVTFGLALIIYNEALKG 89

```

```

RESULT 9
US-09-621-976-7537
; Sequence 7537, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7537
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 90
; OTHER INFORMATION: Xaa = Ala,Gly
US-09-621-976-7537

```

```

Query Match 45.0%; Score 240; DB 4; Length 92;
Best Local Similarity 51.1%; Pred. No. 1.4e-23;
Matches 45; Conservative 19; Mismatches 22; Indels 2; Gaps 1;

QY 1 MACPLEKALDVMTSTFKYSGEGDKFKLNSKELKLTRELPSFLGKRTDEAAFOKLMS 60
DB 1 MACPLDQAIGLLVAIFHKYSGREGDKHTLSKKELKELIKELT--IGSKLQDAEIARLME 58

QY 61 NLDNRDNEVDFOEYCVFLSCIAMMNCNE 88
DB 59 LDRNKDQEVNFQEVYVTFGLALIIYNE 86

```

```

RESULT 10
US-09-051-589-1
; Sequence 1, Application US/09051589
; Patent No. 5930080
; GENERAL INFORMATION:
; APPLICANT: HAGLID, Kenneth G.
; TITLE OF INVENTION: USE OF PROTEIN S-100B IN MEDICINES CONTAINING THE
; TITLE OF INVENTION: PROTEIN S-100B
; FILE REFERENCE: 003300-478
; CURRENT APPLICATION NUMBER: US/09/051,589
; CURRENT FILING DATE: 1998-04-15
; EARLIER APPLICATION NUMBER: SE 9503620-8
; EARLIER FILING DATE: 1995-10-17
; EARLIER APPLICATION NUMBER: PCT/SE96/01305
; EARLIER FILING DATE: 1996-10-15
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Protein S-100b
US-09-051-589-1

```

Query Match 43.3%; Score 231; DB 2; Length 92;

;; NUMBER OF SEQUENCES: 9
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.
;; STREET: 3174 Porter Drive
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; COMPUTER: IBM Compatible
;; OPERATING SYSTEM: DOS
;; SOFTWARE: Fast-SEQ for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/205,680A
;; FILING DATE: Herewith
;; CLASSIFICATION: 530
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Colette C. Muenzen
;; REGISTRATION NUMBER: 39,784
;; REFERENCE/DOCKET NUMBER: PF-0373 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650-855-0555
;; TELEFAX: 650-845-4166
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 92 amino acids
;; TYPE: amino acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: GenBank
;; CLONE: 337730
US-09-205-680A-5

Query Match 43.0%; Score 229; DB 3; Length 92;
Best Local Similarity 48.3%; Pred. No. 3.7e-22;
Matches 42; Conservative 17; Mismatches 28; Indels 0; Gaps 0;
QY 5 LEKALDVNVSTPHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDRAAFQKLSNLD 64
DB 4 LEKAWALIDVPHQISGREGDKHKLSKSELKELINNELSHFLEEIKQEVDVKWETLDN 63
QY 65 NRDNVEVDFQYCVFLSCIAMMCNEFFE 91
DB 64 DGDGECDFQEFMAFVAMVTTACHEFFE 90

RESULT 14
US-09-621-976-7524
; Sequence 7524, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7524
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-7524

Query Match 39.6%; Score 211; DB 4; Length 88;

Best Local Similarity 48.8%; Pred. No. 7.5e-20;
Matches 40; Conservative 18; Mismatches 22; Indels 2; Gaps 1;
QY 11 VNVSTPHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLSNLDNRDNEV 70
DB 9 LLVAIFHKYSGREGDKHTLSKKELKELIKELT--IGSKLQDAETARLMEDLDNRKQDEV 66
QY 71 DFQYCVFLSCIAMMCNEFFE 92
DB 67 NFQYVTFGLGALALIYNEALKG 88

RESULT 15
US-09-621-976-7526
; Sequence 7526, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 7526
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-7526

Query Match 39.6%; Score 211; DB 4; Length 88;
Best Local Similarity 48.8%; Pred. No. 7.5e-20;
Matches 40; Conservative 18; Mismatches 22; Indels 2; Gaps 1;
QY 11 VNVSTPHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLSNLDNRDNEV 70
DB 9 LLVAIFHKYSGREGDKHTLSKKELKELIKELT--IGSKLQDAETARLMEDLDNRKQDEV 66
QY 71 DFQYCVFLSCIAMMCNEFFE 92
DB 67 NFQYVTFGLGALALIYNEALKG 88

Search completed: September 9, 2004, 11:01:10
Job time : 58.6903 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 9, 2004, 10:32:01 ; Search time 212.426 Seconds
(without alignments)
134.340 Million cell updates/sec

Title: US-10-067-618-2

Perfect score: 533

Sequence: 1 MACPLEKALDVMVSTFKYS.....IAMMCNFFEGFDKQPRKK 101

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	533	100.0	101	2	AAR20560 Human mts
2	533	100.0	101	2	AAR80453 Human mts
3	533	100.0	101	3	AAB45534 Human S10
4	533	100.0	101	3	AAB37432 Human mts
5	533	100.0	101	4	AAB72386 Human mts
6	533	100.0	101	6	ABU08513 Human mts
7	533	100.0	101	7	ADD14157 Human src
8	496	93.1	101	4	AAB72387 Marine Mt
9	328	61.5	97	2	AAR27058 Sequence
10	328	61.5	97	2	AAR26406 Sequence
11	323	60.6	97	3	AAB45532 Human S10
12	323	60.6	97	5	ABG96408 Human ova
13	323	60.6	97	7	ABR92148 Human cer
14	323	60.6	97	7	ADB70348 Human calc
15	323	60.6	97	7	ADBS9560 Human Pro
16	323	60.6	97	7	ADBS9556 Human Pro
17	323	60.6	98	6	ABP71986 Human sta
18	323	60.6	98	6	ABU56414 Lung canc
19	323	60.6	98	6	ABU56415 Lung canc
20	323	60.6	98	6	ABU56412 Lung canc
21	248	46.5	90	2	AAY29553 Human lun
22	248	46.5	90	3	AAB44476 Human lun
23	248	46.5	90	3	AAB45536 Human S10
24	248	46.5	90	4	AAB13818 Human lun
25	248	46.5	90	5	ABG96409 Human ova

26	248	46.5	90	6	ABU57620 Different
27	248	46.5	90	7	ADD66508 Human lun
28	248	46.5	90	7	ADE87762 Human lun
29	246	46.2	105	4	AAG73796 Human col
30	246	46.2	105	5	ABP41087 Human ova
31	244	45.8	90	3	AAY57068 Calbindin
32	244	45.8	94	3	AAB45531 Human S10
33	244	45.8	94	4	AAM40258 Human ova
34	244	45.8	94	5	ABG96406 Human ova
35	241	45.2	93	7	ADE61764 Human Pro
36	239	44.8	93	7	ADE61762 Rat Prote
37	233.5	43.8	110	3	AAB45535 Human S10
38	231	43.3	97	5	ABP51390 Human MDD
39	231	43.3	97	6	ABU11560 Human MDD
40	229	43.0	92	5	ABB97495 Novel hum
41	229	43.0	92	6	ABR58703 Human Can
42	229	43.0	92	6	ABU63333 Human S10
43	229	43.0	92	8	ADE77019 Human pro
44	229	43.0	101	3	AAB45533 Human S10
45	229	43.0	124	7	ADE08957 Novel pro

ALIGNMENTS

RESULT 1

AAR20560
ID AAR20560 standard; protein; 101 AA.

XX AC AAR20560;

XX DT 25-MAR-2003 (revised)

XX DT 14-MAY-1992 (first entry)

XX DE Human mts protein.

XX KW Metastatic cancer; antibodies; mouse; lung; liver; kidney; thyroid;

XX KW breast cancer; cell growth.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Protein 1..101

FT Peptide /label= mts-1

FT Peptide 2..11

FT Peptide /note= "antigenic"

FT Peptide 22..37

FT Peptide /note= "antigenic; calcium binding domain"

FT Peptide 42..54

FT Peptide /note= "antigenic"

FT Peptide 87..101

FT Peptide /note= "antigenic"

XX WO9200757-A.

XX PD 23-JAN-1992.

XX PF Aar26406 Sequence

XX PF Aab45532 Human S10

XX PF Abg96408 Human ova

XX PR AbR92148 Human cer

XX PR Adb70348 S100 calc

XX PA Ades9560 Human Pro

XX PI Ades9556 Human Pro

XX PI Zain S, Lukanidin E;

XX DR Abp71986 Human sta

XX DR Abus6414 Lung canc

XX DR Abus6415 Lung canc

XX XX Abus6412 Lung canc

XX PT Aay29553 Human lun

XX PT Aab44476 Human lun

XX PT Aab45536 Human S10

XX PT Aae13818 Human lun

XX PS Abg96409 Human ova

XX PS Claim 9; Fig 2; 82pp; English.

Metastatic cancer diagnosis by detection of mts-1 gene or protein - using antibody treatment of cancer and tumours of e.g. kidney, thyroid, lung and liver.

XX The sequence was deduced from the DNA sequence obtd. by screening a human
CC cDNA library with mouse mts-1 cDNA probes. The antigenic mts-1 peptides
CC (see features) derived from the protein and anti- bodies raised to them
CC are useful in the diagnosis of metastatic cancer, e.g. lung, kidney,
CC thyroid or breast cancer. The peptide comprising the calcium binding site
CC generates antibodies reactive with many members of the calcium binding
CC protein family; the other three peptides are unique to mts-1 and generate
CC antibodies specific only for this protein. Cell lines capable of
CC expressing mts-1 are useful as model systems for in vitro and in vivo
CC anti-metastasis drug screening. Pharmaceutical compns. contg. the mts-1
CC protein or anti-cancer reagents may be used to promote cell growth, or
CC for treating cancer, respectively. (Updated on 25-MAR-2003 to correct PA
CC field.)
XX SQ

Sequence 101 AA;

Query Match 100.0%; Score 533; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. NO. 3.9e-59;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACPLEKALDVWVSTFHKYSKGEKDFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLS 60
DB 1 MACPLEKALDVWVSTFHKYSKGEKDFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLS 60
QY 61 NLDNRDNEVDFOEYCVFLSCIAMMCMNEFFEGFPDKQPRKK 101
DB 61 NLDNRDNEVDFOEYCVFLSCIAMMCMNEFFEGFPDKQPRKK 101

RESULT 2

AAR80453
ID AAR80453 standard; protein; 101 AA.

AC AAR80453;

XX 27-DEC-1995 (first entry)

DE Human mts-1 protein.

XX Metastasis; cancer; mts-1 gene; tumour; therapy.

XX Homo sapiens.

XX WO9520656-A1.

XX 03-AUG-1995.

XX 31-JAN-1995; 95WO-US001214.

XX 31-JAN-1994; 94US-00190560.

XX (RESE) RESEARCH CORP TECHNOLOGIES INC.

XX Zain S, Lukanidin E;

XX WPI; 1995-275441/36.

XX N-PSDB; AAQ99177.

XX Nucleic acid encoding human mts-1, antigenic fragments and antibodies -
PT useful for diagnosis of malignant cancer and metastatic potential of
PT tumour cells.

XX Claim 12; Page 92; 124pp; English.

XX A human cDNA library was constructed in lambda-gt10 using poly(A)+ RNA
CC prep'd. from HeLa cells. The library was screened with a 32P- labeled
CC mouse mts-1 cDNA probe. A clone was obtd. which comprised the full-length
CC human mts-1 gene. The encoded protein is used for the diagnosis or
CC therapy of cancer, and to raise antibodies
XX SQ

Sequence 101 AA;

Query Match 100.0%; Score 533; DB 2; Length 101;
Best Local Similarity 100.0%; Pred. NO. 3.9e-59;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACPLEKALDVWVSTFHKYSKGEKDFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLS 60
DB 1 MACPLEKALDVWVSTFHKYSKGEKDFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLS 60
QY 61 NLDNRDNEVDFOEYCVFLSCIAMMCMNEFFEGFPDKQPRKK 101
DB 61 NLDNRDNEVDFOEYCVFLSCIAMMCMNEFFEGFPDKQPRKK 101

RESULT 3

AAB45534
ID AAB45534 standard; protein; 101 AA.

AC AAB45534;

XX 22-FEB-2001 (first entry)

XX Human S100A4 protein.

XX S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;
KW calcium-binding protein; calcium homeostasis; cardiac muscle;
KW pumping capacity; myocardial cell; systolic calcium ion release;
KW sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;
KW valve defect.

XX Homo sapiens.

XX DE19915485-A1.

XX 19-OCT-2000.

XX 07-APR-1999; 99DE-01015485.

XX 07-APR-1999; 99DE-01015485.

XX (KATU/) KATUS H A.

XX (REMP/) REMPPIS A.

XX Katus HA, Remppis A;

XX WPI; 2000-673510/66.

XX N-PSDB; AAC81804.

XX Composition containing S100 protein, corresponding nucleic acid or
PT vector, useful for treating cardiomyopathy and cardiac insufficiency.

XX Claim 35; Page 11; 36pp; German.

XX This invention describes a novel composition for treating primary or
CC secondary cardiomyopathy or cardiac insufficiency contains at least one
CC S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or
CC fragments, or a gene transfer vector containing (II), optionally
CC formulated with auxiliaries and/or carriers. (I) are calcium-binding
CC proteins involved in calcium homeostasis, so their overexpression in
CC cardiac muscle will improve pumping capacity (and overall capacity) of
CC the heart. In cultured myocardial cells they increase the contraction and
CC relaxation rates associated with increased systolic calcium ion release
CC from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are
CC used to treat cardiomyopathy (CMP) where inherited or caused by
CC spontaneous mutations and ischemic CMP caused by arteriosclerosis,
CC dilative CMP caused by toxic/infectious disease, cardiac disease caused
CC by pulmonary and/or arterial hypertension, and structural disease caused
CC by rhythm disorders or valve defects, generally any condition associated
CC with reduced contractile force. Unlike calmodulin, which is expressed
CC ubiquitously, (I) show tissue-specific expression and treat the
CC underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac
CC disease
XX SQ

Sequence 101 AA;

Query Match 100.0%; Score 533; DB 3; Length 101;
 Best Local Similarity 100.0%; Pred. No. 3.9e-59;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
 Db 1 MACPLEKALDVMVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
 QY 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101
 Db 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101

RESULT 4
 AAB37432
 ID AAB37432 standard; protein; 101 AA.
 XX
 AC AAB37432;
 XX
 DT 21-FEB-2001 (first entry)
 XX
 DE Human mts-1.
 XX
 KW Human; mts-1; cytostatic; cancer; metastasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200064475-A1.
 XX
 XX 02-NOV-2000.
 XX
 PF 20-APR-2000; 2000WO-US011006.
 XX
 PR 23-APR-1999; 99US-00298625.
 XX
 PA (RESE) RESEARCH CORP TECHNOLOGIES INC.
 XX
 PI Lukanidin E;
 XX
 DR WPI; 2000-687266/67.
 XX
 DR N-PSDB; AAC68131, AAC68132.
 XX
 PT Treating cancer and preventing metastasis comprises administration of an
 PT anti-mts-1 protein antibody or antisense oligonucleotide.
 XX
 PS Example 6; Fig 2; 155pp; English.
 XX
 CC The present sequence is human mts-1. Mts-1 protein is a calcium-binding
 CC protein, and is thought to have a role in myoepithelial cell
 CC differentiation. The present invention relates to methods for treating
 CC cancer and preventing metastases, comprising the administration of a
 CC composition directed against the mts-1 protein
 XX
 SQ Sequence 101 AA;

Query Match 100.0%; Score 533; DB 3; Length 101;
 Best Local Similarity 100.0%; Pred. No. 3.9e-59;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
 Db 1 MACPLEKALDVMVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
 QY 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101
 Db 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101

RESULT 5
 AAB72386
 ID AAB72386 standard; protein; 101 AA.
 XX

AC AAB72386;
 XX 24-MAY-2001 (first entry)
 DT Human Mts 1 protein amino acid sequence.
 DE
 XX Mts 1; Mts1-del75; Mts1-4S; S100A4; neuronal cell growth; cardiac arrest;
 KW neuronal degeneration; Parkinson's disease; Alzheimer's disease;
 KW Down's syndrome; stroke; sciatic crush; spinal cord injury; human.
 XX
 OS Homo sapiens.
 XX
 PN WO200118043-A2.
 XX
 PD 15-MAR-2001.
 XX
 PF 07-SEP-2000; 2000WO-US024495.
 XX
 PR 10-SEP-1999; 99US-00393433.
 XX
 PA (PROL-) PROLIFIA INC.
 XX
 PI Bock E, Lukanidin EM, Berezin V;
 XX
 DR WPI; 2001-235188/24.
 XX
 XX New isolated functional derivatives of Mts protein for stimulating
 PT neuronal cell growth to treat neurological conditions e.g. Parkinson's
 PT disease, Alzheimer's disease, Down's syndrome and stroke.
 XX
 PS Disclosure; Page; 60pp; English.
 XX
 CC This invention relates to functional derivatives of an Mts protein,
 CC particularly Mts1-del75, and Mts1-4S. Mts 1 (also known as S100A4) is a
 CC member of the S100 gene family. Mts1 proteins and their functional
 CC derivatives are used to stimulate the growth of neuronal cells and in the
 CC treatment of neurological conditions including those characterised by
 CC neuronal degeneration, death or injury such as Parkinson's disease,
 CC Alzheimer's disease, Down's syndrome, stroke, cardiac arrest, sciatic
 CC crush, spinal cord injury, injury to sensory neurons or degenerative
 CC disease of the retina. The present sequence represents the human Mts 1
 CC protein, from which the Mts proteins of the invention may be derived.
 CC NOTE: The present sequence is not shown in the specification but it has
 CC been derived from the sequence given in US patent 5801142 (referred to on
 CC page 10 of the specification)
 XX
 SQ Sequence 101 AA;

Query Match 100.0%; Score 533; DB 4; Length 101;
 Best Local Similarity 100.0%; Pred. No. 3.9e-59;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
 Db 1 MACPLEKALDVMVSTFHKYSKGGDKFKLNKSELKELLTRELPSFLGKRTDEAAAFQKLMS 60
 QY 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101
 Db 61 NLDNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDKQPRKK 101

RESULT 6
 ABU08513
 ID ABU08513 standard; protein; 101 AA.
 XX
 AC ABU08513;
 XX
 DT 27-MAY-2003 (first entry)
 XX
 DE Human mts-1 protein.
 XX
 KW Human; mts-1; metastatic cancer; p53; cytostatic; gene therapy.
 XX

OS Homo sapiens.
XX US2002172680-A1.
XX 21-NOV-2002.
XX 29-APR-2002; 2002US-00135152.
XX 09-JUL-1990; 90US-00550600.
XX 25-NOV-1992; 92US-00981455.
XX 31-JAN-1994; 94US-00190560.
XX 06-JUN-1995; 95US-00468942.
XX 23-APR-1999; 99US-00298635.
XX (LUKA/) LUKANIDIN E.
XX Lukanidin E;
XX WPI; 2003-328422/31.
XX N-PSDB; ABX93578.
XX Treating cancer by administering a reagent directed against the mts-1
XX protein or an oligonucleotide capable of binding to mts-1 mRNA.
XX Example 6; Fig 2; 64pp; English.
XX The invention relates to treating cancer comprising administering a
XX reagent directed against the mts-1 (not defined) protein or an
XX oligonucleotide capable of binding to mts-1 mRNA. Also include are a
XX method of inactivating, destroying or nullifying a mts-1 protein or cells
XX and a method of inhibiting metastasis in a cancerous cell. Experiments
XX showed that mts-1 inhibited the phosphorylation of full-size p53 and the
XX C-terminal protein fragment by PKC. Addition of the same concentrations
XX of mts-1 to the PKC reaction mixture did not affect the phosphorylation
XX of the N-terminal and DNA-binding domains of p53. The method is useful
XX for treating cancers associated with the mts-1 gene. The present sequence
XX represents human mts-1 protein
XX
SQ Sequence 101 AA;
Query Match 100.0%; Score 533; DB 6; Length 101;
Best Local Similarity 100.0%; Pred. No. 3.9e-59;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MACPLEKALDVMTSTFKYSGEGDKFKLNKSKELLLTRELPSFLGKRTDEAAFOKLMS 60
Db 1 MACPLEKALDVMTSTFKYSGEGDKFKLNKSKELLLTRELPSFLGKRTDEAAFOKLMS 60
QY 61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101
Db 61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101
RESULT 7
ID ADD14157
XX ADD14157 standard; protein; 101 AA.
XX ADD14157;
XX 01-JAN-2004 (first entry)
XX Human src biomarker polypeptide SEQ ID NO:346.
XX predictor set; protein tyrosine kinase activity modulator;
XX protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
XX gene therapy; drug sensitivity; genetic profile; cancer; human.
XX Homo sapiens.
XX WO2003062395-A2.
XX 31-JUL-2003.
XX

PF 17-JAN-2003; 2003WO-US001981.
XX
XX 18-JAN-2002; 2002US-0350061P.
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX Huang F, Fairchild CR, Lee FY, Shaw P;
XX WPI; 2003-636735/60.
XX N-PSDB; ADD14760.
XX New polynucleotides and polypeptides for predicting the activity of
XX compounds that interact with protein tyrosine kinases and/or protein
XX tyrosine kinase pathways.
XX Claim 10; SEQ ID NO 346; 139pp; English.
XX The present invention describes a predictor set comprising a plurality of
XX polynucleotides or polypeptides whose expression pattern is predictive of
XX the response of cells to treatment with a compound that modulates protein
XX tyrosine kinase activity or members of the protein tyrosine kinase
XX pathway. Also described: (1) predicting whether a compound is capable of
XX modulating the activity of cells, comprising obtaining a sample of cells,
XX determining whether the cells express a plurality of markers, and
XX correlating the expression of the markers to the compound's ability to
XX modulate the activity of the cells; (2) a plurality of cell lines for
XX identifying polynucleotides and polypeptides whose expression levels
XX correlate with compound sensitivity or resistance of cells associated
XX with a disease state; and (3) identifying polynucleotides and
XX polypeptides that predict compound sensitivity or resistance of cells
XX associated with a disease state, comprising subjecting the plurality of
XX cell lines to one or more compounds, analysing the expression pattern of
XX a microarray of polynucleotides or polypeptides, and selecting
XX polynucleotides or polypeptides that predict the sensitivity or
XX resistance of cells associated with a disease state by using the
XX expression pattern of the microarray. The polynucleotides and
XX polypeptides have cytostatic activities, and can be used in gene therapy.
XX The polynucleotides and polypeptides are useful in predicting the
XX activity of compounds that interact with protein tyrosine kinases and/or
XX protein tyrosine kinase pathways. These may be used in determining drug
XX sensitivity in patients to allow the development of individualized
XX genetic profiles which aid in treating diseases and disorders (e.g.
XX cancer) based on patient response at a molecular level. The present
XX sequence is used in the exemplification of the present invention.
XX
SQ Sequence 101 AA;
Query Match 100.0%; Score 533; DB 7; Length 101;
Best Local Similarity 100.0%; Pred. No. 3.9e-59;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MACPLEKALDVMTSTFKYSGEGDKFKLNKSKELLLTRELPSFLGKRTDEAAFOKLMS 60
Db 1 MACPLEKALDVMTSTFKYSGEGDKFKLNKSKELLLTRELPSFLGKRTDEAAFOKLMS 60
QY 61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101
Db 61 NLDNSNRDNEVDFOEYCVFLSCIAMMCNEFFEGFPDQPRKK 101
RESULT 8
ID AAB72387
XX AAB72387 standard; protein; 101 AA.
XX AAB72387;
XX 24-MAY-2001 (first entry)
XX Murine Mts 1 protein amino acid sequence.
XX Mts 1; Mts1-del75; Mts1-4S; S100A4; neuronal cell growth; cardiac arrest;
XX neuronal degeneration; Parkinson's disease; Alzheimer's disease;
XX Down's syndrome; stroke; sciatic crush; spinal cord injury; mouse.
XX

XX OS Mus sp.
XX PN WO200118043-A2.
XX PD 15-MAR-2001.
XX PF 07-SEP-2000; 2000WO-US024495.
XX PR 10-SEP-1999; 99US-00393433.
XX PA (PROL-) PROLIFIA INC.
XX PI Bock E, Lukanidin EM, Berezin V;
XX WIPI; 2001-235188/24.
XX DR New isolated functional derivatives of Mts protein for stimulating
XX PT neuronal cell growth to treat neurological conditions e.g. Parkinson's
XX PT disease, Alzheimer's disease, Down's syndrome and stroke.
XX PS Disclosure; Page; 60pp; English.
XX CC This invention relates to functional derivatives of an Mts protein,
XX particularly Mts1-del75, and Mts1-45. Mts 1 (also known as S100A4) is a
XX member of the S100 gene family. Mts1 proteins and their functional
XX derivatives are used to stimulate the growth of neuronal cells and in the
XX treatment of neurological conditions including those characterised by
XX neuronal degeneration, death or injury such as Parkinson's disease,
XX Alzheimer's disease, Down's syndrome, stroke, cardiac arrest, sciatic
XX crush, spinal cord injury, injury to sensory neurons or degenerative
XX disease of the retina. The present sequence represents the murine Mts 1
XX protein, from which the Mts proteins of the invention may be derived.
XX NOTE: The present sequence is not shown in the specification but it has
XX been derived from the sequence given in Genes Dev. 3, 1989 (referred to
XX on page 10 of the specification)
XX SQ Sequence 101 AA;

Query Match 93.1%; Score 496; DB 4; Length 101;
Best Local Similarity 93.1%; Pred. No. 1.9e-54;
Matches 94; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMVSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db 1 MARPLEALDVIIVSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKVM 60

QY 61 NLDNSNRDNEVDFOYCVFLSCIAMMCNEFFEGPDKQPRKK 101
Db 61 NLDNSNRDNEVDFOYCVFLSCIAMMCNEFFEGPDKQPRKK 101

RESULT 9
AAR27058
ID AAR27058 standard; protein; 97 AA.
XX AC AAR27058;
XX DT 25-MAR-2003 (revised)
XX DT 25-FEB-1993 (first entry)
XX DE Sequence of small Ca++ binding proteins encoded by Can19 clone 19.
XX KW Can19; tumour suppressor gene; cancer; therapy.
XX OS Homo sapiens.
XX PN WO9215602-A1.
XX PD 17-SEP-1992.
XX PF 28-FEB-1992; 92WO-US001624.
XX XX

PR 28-FEB-1991; 91US-00662216.
XX (DAND) DANA FARBER CANCER INST INC.
XX Sager R;
XX WIPI; 1992-331663/40.
XX DR N-PSDB; AAQ28760.
XX PT Diagnosis and treatment of cancer - using candidate tumour suppressor
XX genes or the corresp. antibodies.
XX PS Disclosure; Page 31; 54pp; English.
XX CC A clone originally termed clone 19, and now referred to as Can19,
XX represents a gene expressed in normal mammary epithelial cell strains but
XX not in tumor-derived cell lines. Sequence comparisons have shown that
XX Can19 is a member of the S100 gene family, encoding small Ca++ binding
XX proteins (about 10 kD) with diverse functions. These proteins have two
XX "EF hands", domains where Ca2+ is bound. Can19 is also related in
XX structure to the small regulatory subunit of calpactin, p11. Can19 is not
XX expressed in breast tumor cells. Can19 appears to be negatively regulated
XX in tumors, in contrast to calyculin. (Updated on 25-MAR-2003 to correct
XX PN field.) (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 97 AA;

Query Match 61.5%; Score 328; DB 2; Length 97;
Best Local Similarity 62.5%; Pred. No. 3.1e-33;
Matches 60; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMVSTFKYSGEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db 1 MCSSLFQALAVLTTFHKYSCQEGDKFKLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60

QY 61 NLDNSNRDNEVDFOYCVFLSCIAMMCNEFFEGPDK 96
Db 61 NLDNSNRDNEVDFOYCVFLSCIAMMCNEFFEGPDK 96

RESULT 10
AAR26406
ID AAR26406 standard; protein; 97 AA.
XX AC AAR26406;
XX DT 25-MAR-2003 (revised)
XX DT 27-FEB-1993 (first entry)
XX DE Sequence of the clone 19 gene product.
XX KW Clone 19; diagnosis; prognosis; cancer; tumour.
XX OS Homo sapiens.
XX PN WO9215600-A1.
XX PD 17-SEP-1992.
XX PF 28-FEB-1992; 92WO-US001625.
XX XX 28-FEB-1991; 91US-00662198.
XX PA (DAND) DANA FARBER CANCER INST INC.
XX PI Sager R, Lee SW, Tomasetto C;
XX XX WIPI; 1992-331662/40.
XX DR N-PSDB; AAQ28007.
XX PT Clone 19 gene prodn. and its DNA and antibody - for diagnosis, prognosis
XX and treatment of solid tumours, especially breast cancer.
XX XX

PS Disclosure; Page 12-13; 20pp; English.

CC Clone 19 was derived from normal human mammary epithelial cells strain

CC 76N. Clone 19 represents a gene expressed in normal mammary epithelial

CC cell strains but not in tumour-derived cell lines. Sequence comparisons

CC have shown that it is a member of the S100 gene family, encoding small

CC Ca++ binding proteins (about 10kd) with diverse functions. These proteins

CC have two 'EF hands', domains where Ca++ is bound, in contrast to calmodulin

CC proteins which have four. Clone 19 is also related in structure to the

CC small regulatory subunit of calpactin, p11. MRP8 and MRP14 are also

CC related, and are S100 proteins expressed by macrophages during chronic

CC inflammation. (Updated on 25-MAR-2003 to correct FN field.) (Updated on

CC 25-MAR-2003 to correct PA field.)

XX Sequence 97 AA;

XX SQ

Query Match 61.5%; Score 328; DB 2; Length 97;

Best Local Similarity 62.5%; Pred. No. 3.1e-33;

Matches 60; Conservative 17; Mismatches 19; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTSTFKYSGKGDGKFKLNSKELKELLTRELPSFLGKRTDEAAFOKLMS 60

DB 1 MCSLEQALAVLVTTTFHKYSCQEGDKFKLSKGMKELLHKELPSFVGKVDSEGLKLMG 60

QY 61 NLDNSRDNEVDFOEYCVFLSCIAMMNCNEFFEGFPDK 96

DB 61 NLDNSDQQVDFQYAVFLALITVNCNDFFOGCPDR 96

RESULT 11

AAB45532

ID AAB45532 standard; protein; 97 AA.

XX AAB45532;

XX 22-FEB-2001 (first entry)

XX Human S100A2 protein.

XX S100 protein; human; treatment; cardiomyopathy; cardiac insufficiency;

XX calcium-binding protein; calcium homeostasis; cardiac muscle;

XX pumping capacity; myocardial cell; systolic calcium ion release;

XX sarcoplasmic reticulum; cardiac disease; hypertension; rhythm disorder;

XX valve defect.

XX Homo sapiens.

XX DE19915485-A1.

XX 19-OCT-2000.

XX 07-APR-1999; 99DE-01015485.

XX 07-APR-1999; 99DE-01015485.

XX (KATU/) KATUS H A.

XX (REMP/) REMPFIS A.

XX Katus HA, Remppis A;

XX WPI; 2000-673510/66.

XX Composition containing S100 protein, corresponding nucleic acid or

XX vector, useful for treating cardiomyopathy and cardiac insufficiency.

XX Claim 35; Page 9; 36pp; German.

XX This invention describes a novel composition for treating primary or

XX secondary cardiomyopathy or cardiac insufficiency contains at least one

XX S100 protein (I) or nucleic acid (II) encoding (I), or their mutants or

XX fragments, or a gene transfer vector containing (II), optionally

XX formulated with auxiliaries and/or carriers. (I) are calcium-binding

XX proteins involved in calcium homeostasis, so their overexpression in

CC cardiac muscle will improve pumping capacity (and overall capacity) of

CC the heart. In cultured myocardial cells they increase the contraction and

CC relaxation rates associated with increased systolic calcium ion release

CC from the sarcoplasmic reticulum (SR) and calcium re-uptake by SR. (I) are

CC used to treat cardiomyopathy (CMP) where inherited or caused by

CC spontaneous mutations and ischemic CMP caused by arteriosclerosis,

CC dilative CMP caused by toxic/infectious disease, cardiac disease caused

CC by pulmonary and/or arterial hypertension, and structural disease caused

CC by rhythm disorders or valve defects, generally any condition associated

CC with reduced contractile force. Unlike calmodulin, which is expressed

CC ubiquitously, (I) show tissue-specific expression and treat the

CC underlying defect in the sarcoplasmic reticulum (SR) that causes cardiac

CC disease

XX Sequence 97 AA;

XX SQ

Query Match 60.6%; Score 323; DB 3; Length 97;

Best Local Similarity 61.5%; Pred. No. 1.3e-32;

Matches 59; Conservative 18; Mismatches 19; Indels 0; Gaps 0;

QY 1 MACPLEKALDVMTSTFKYSGKGDGKFKLNSKELKELLTRELPSFLGKRTDEAAFOKLMS 60

DB 1 MCSLEQALAVLVTTTFHKYSCQEGDKFKLSKGMKELLHKELPSFVGKVDSEGLKLMG 60

QY 61 NLDNSRDNEVDFOEYCVFLSCIAMMNCNEFFEGFPDK 96

DB 61 NLDNSDQQVDFQYAVFLALITVNCNDFFOGCPDR 96

RESULT 12

ABG96408

ID ABG96408 standard; protein; 97 AA.

XX ABG96408;

XX 11-DEC-2002 (first entry)

XX Human ovarian cancer marker M68.

XX Human; ovarian cancer; marker; cancer; familial history; brain disorder;

XX central nervous system disorder; bacterial meningitis; viral meningitis;

XX Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;

XX brain herniation; inflammation; encephalitis; testicular disorder;

XX nontuberculous granulomatous orchitis; connective tissue disorder;

XX heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;

XX histological type; carcinogenic; ovarian cancer marker.

XX Homo sapiens.

XX WO200271928-A2.

XX 19-SEP-2002.

XX 14-MAR-2002; 2002WO-US007826.

XX 14-MAR-2001; 2001US-0276025P.

XX 14-MAR-2001; 2001US-0276026P.

XX 10-AUG-2001; 2001US-0311732P.

XX 19-SEP-2001; 2001US-0323580P.

XX 26-SEP-2001; 2001US-0324967P.

XX 26-SEP-2001; 2001US-0325102P.

XX 26-SEP-2001; 2001US-0325149P.

XX (MILL-) MILLENNIUM PHARM INC.

XX Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;

XX Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;

XX Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;

XX WPI; 2002-723277/78.

XX N-PSDB; ABS76507.

XX Assessing whether a patient is afflicted with ovarian cancer, useful in

PT assessing the stage or progression of the disease, comprises comparing
PT the expression level of a cancer marker in a sample from a patient and
XX from a non cancer patient.

PS Disclosure; Page 392; 481pp; English.

XX The present invention relates to a new method for assessing whether a
CC patient is afflicted with ovarian cancer. The method involves comparing
CC the expression level of a marker in a patient sample and the normal level
CC of expression of the marker in a control non-ovarian cancer sample, where
CC the marker is selected from 363 cancer markers described in the
CC specification. The method of the invention is useful in diagnosing or
CC characterising cancer, in detecting the presence of cancer as early as
CC possible, and the recurrence of ovarian cancer. The method may also be of
CC particular use with patients having an enhanced risk of developing
CC ovarian cancer (e.g. patients having a familial history of ovarian
CC cancer). The cancer markers may be used in the management and treatment
CC of e.g. brain and central nervous system disorders (e.g. bacterial and
CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
CC testicular disorders (e.g. nontuberculous granulomatous orchitis),
CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
CC disease or atherosclerosis). The compositions and methods may also be
CC used in assessing the histological type of neoplasm associated with
CC ovarian cancer, monitoring the progression of ovarian cancer, determining
CC whether ovarian cancer has metastasized or is likely to metastasize,
CC selecting a composition for inhibiting ovarian cancer, assessing the
CC ovarian carcinogenic potential of a compound, or inhibiting ovarian
CC cancer or at risk of developing ovarian cancer. The present amino acid
CC sequence represents one of the ovarian cancer markers described in the
CC invention

XX SQ Sequence 97 AA;

Query Match 60.6%; Score 323; DB 5; Length 97;
Best Local Similarity 61.5%; Pred. No. 1.3e-32;
Matches 59; Conservative 18; Mismatches 19; Indels 0; Gaps 0;
QY 1 MACPLEKALDVMVSTFKYSGKGDKFKLNKSELKELLTRELPSFLGKRTDEAAFKLMS 60
Db 1 MCSLSLEQALAVLVTTHFKYSCQGDGKFKLSKGMKELLHKELPSFVGKVDGLKKLNG 60
QY 61 NLDNSRNEVDFOEYCVFLSCIAMMCNEPFGPDK 96
Db 61 SLDNSDQQVDFQEYAVFLALITVMCNDFFQGCPCR 96

RESULT 13
ABR92148
ID ABR92148 standard; protein; 97 AA.
XX ABR92148;
AC ABR92148;
XX 10-SEP-2003 (first entry)
XX Human cervical cancer cell marker protein SEQ ID NO:206.
DE Human, cervical cancer; cervical cancer marker; cancer therapy;
XX detection; gene therapy; vaccine.
XX Homo sapiens.
XX WO2002101075-A2.
XX 19-DEC-2002.
XX 12-JUN-2002; 2002WO-US018638.
XX 13-JUN-2001; 2001US-0298155P.
XX 13-JUN-2001; 2001US-0298159P.
XX 14-NOV-2001; 2001US-0335936P.

PA (MILL-) MILLENNIUM PHARM INC.

XX Schlegel R, Chen Y, Zhao Y, Monahan JE, Kamatkar S;
PI Gannavarapu M, Glatt K, Hoersch S;

XX WPI; 2003-156967/15.

DR N-PSDB; ACFI2931.

XX New isolated nucleic acid molecule useful for detecting, characterizing,
PT preventing and treating human cervical cancers, in various prognostic and
PT diagnostic assays, in pharmacogenomics and in monitoring clinical trials.
XX Claim 4; Page 361-362; 386pp; English.

XX ACFI2828 to ACFI2947 encode the human cervical cancer marker proteins (I)
CC given in ABR92047 to ABR92184. A higher level of expression of (I) than
CC normal indicates the presence of cervical cancer. Also described: (1) a
CC vector (II) containing (I); (2) a host cell (III) containing (I); and (3)
CC assessing (M1) whether a patient is afflicted with cervical cancer,
CC comprising comparing the level of expression of a marker in a patient's
CC sample, and the normal level of expression of the marker in a control non
CC -cervical cancer sample, where a significant increase in the level of
CC expression of the marker in the patient's sample relative to that in the
CC control sample is an indication that the patient is afflicted with
CC cervical cancer. (I) has cytostatic activity, and can be used in gene
CC therapy and in vaccines. (I) is useful in detecting, characterising,
CC preventing and treating human cervical cancers. (I) may also be used in
CC various prognostic and diagnostic assays, pharmacogenomics and in
CC monitoring clinical trials

XX SQ Sequence 97 AA;

Query Match 60.6%; Score 323; DB 6; Length 97;
Best Local Similarity 61.5%; Pred. No. 1.3e-32;
Matches 59; Conservative 18; Mismatches 19; Indels 0; Gaps 0;
QY 1 MACPLEKALDVMVSTFKYSGKGDKFKLNKSELKELLTRELPSFLGKRTDEAAFKLMS 60
Db 1 MCSLSLEQALAVLVTTHFKYSCQGDGKFKLSKGMKELLHKELPSFVGKVDGLKKLNG 60
QY 61 NLDNSRNEVDFOEYCVFLSCIAMMCNEPFGPDK 96
Db 61 SLDNSDQQVDFQEYAVFLALITVMCNDFFQGCPCR 96

RESULT 14
ADB70348
ID ADB70348 standard; protein; 97 AA.
XX ADB70348;
AC ADB70348;
XX 04-DEC-2003 (first entry)
XX S100 calcium-binding protein A2 SEQ ID NO:40.

XX cancer; malignant pleural mesothelioma; MPW; lung adenocarcinoma;
XX squamous carcinoma; medulloblastoma; prostate cancer; breast cancer;
XX diffuse large B-cell lymphoma; follicular lymphoma; ovarian cancer;
XX human.

XX Homo sapiens.

XX WO2003021229-A2.

XX 13-MAR-2003.

XX 05-SEP-2002; 2002WO-US028203.

XX 05-SEP-2001; 2001US-0317389P.

XX 30-AUG-2002; 2002US-00236031.

XX (BGHM) BRIGHAM & WOMENS HOSPITAL INC.

```

PI Gordon GJ, Jensen RV, Gullans SR, Bueno R;
XX WPI; 2003-290233/28.
DR N-PSDB; ADB70347.
XX
XX Diagnosing cancer cells in tissue sample, or determining prognosis or
PT outcome of cancer patient, by calculating ratio of expression levels of
PT genes that are differentially expressed in cancer and non cancer tissues.
XX
XX Claim 77; Page 161-162; 396pp; English.
XX
XX The present invention describes a method (M1) for diagnosing the presence
CC of cancer cells or non-cancer cells in a tissue sample, or determining
CC the prognosis or outcome of a cancer patient. M1 involves providing a set
CC of genes that are differentially expressed in cancerous or non-cancerous
CC conditions, determining the expression levels of the set of genes and
CC calculating a ratio of the expression levels of the differentially
CC expressed genes. M1 is useful for diagnosing the presence of cancer cells
CC or non-cancer cells in a tissue sample, where the cancer is malignant
CC pleural mesothelioma (MPM), lung adenocarcinoma, squamous carcinoma,
CC medulloblastoma, prostate cancer, breast cancer, diffuse large B-cell
CC lymphoma, follicular lymphoma and ovarian cancer, and for determining
CC prognosis or outcome of a cancer patient. The ratio of expression levels
CC of differentially expressed genes is used as an indicator of cancer type,
CC cancer class, and/or cancer prognosis, all of which are useful for
CC determining a course of treatment of a patient. The present sequence
CC represents a human protein which is used in an example from the present
XX invention.
XX
XX Sequence 97 AA;
SQ
Query Match 60.6%; Score 323; DB 7; Length 97;
Best Local Similarity 61.5%; Pred. No. 1.3e-32;
Matches 59; Conservative 18; Mismatches 19; Indels 0; Gaps 0;
QY 1 MACPLEKALDVMVSTFHKYSCKEGDKFLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db 1 MCSLEQALAVLVTTFHFKYSQCEGDKFLSKGEMKELLHKELPSFVGKEVDEGLKLMG 60
QY 61 NLDSNRDNEVDFOEYCVFLSCIAMWCNEFFFGFDPK 96
Db 61 SLDENSQQVDFOEYAVFLALITWCNDFFGCCPDR 96
RESULT 15
ADE59560
ID ADE59560 standard; protein; 97 AA.
XX
XX ADE59560;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human Protein P29034, SEQ ID NO 5456.
XX
XX Human; pain; neuronal tissue; gene therapy;
XX spinal segmental nerve injury; chronic constriction injury; CCI;
XX spared nerve injury; SNI; Chung.
XX
XX Homo sapiens.
XX
XX WO2003016475-A2.
XX
XX 27-FEB-2003.
XX
XX 14-AUG-2002; 2002WO-US025765.
XX
XX 14-AUG-2001; 2001US-0312147P.
XX
XX 01-NOV-2001; 2001US-0346382P.
XX
XX 26-NOV-2001; 2001US-0333347P.
XX
XX (GEO) GEN HOSPITAL CORP.
XX (FARB) BAYER AG.
XX

```

```

PI Woolf C, D'urso D, Befort K, Costigan M;
XX WPI; 2003-268312/26.
DR GENBANK; P29034.
XX
XX New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
PT
XX Claim 1; Page; 1017pp; English.
XX
XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 97 AA;
SQ
Query Match 60.6%; Score 323; DB 7; Length 97;
Best Local Similarity 61.5%; Pred. No. 1.3e-32;
Matches 59; Conservative 18; Mismatches 19; Indels 0; Gaps 0;
QY 1 MACPLEKALDVMVSTFHKYSCKEGDKFLNKSELKELLTRELPSFLGKRTDEAAFOKLMS 60
Db 1 MCSLEQALAVLVTTFHFKYSQCEGDKFLSKGEMKELLHKELPSFVGKEVDEGLKLMG 60
QY 61 NLDSNRDNEVDFOEYCVFLSCIAMWCNEFFFGFDPK 96
Db 61 SLDENSQQVDFOEYAVFLALITWCNDFFGCCPDR 96
Search completed: September 9, 2004, 10:53:00
Job time : 216.426 secs

```